Department of Production Animal Medicine
Faculty of Veterinary Medicine
University of Helsinki
Finland

INHERITED DEVELOPMENTAL DISEASES RELATED TO REPRODUCTIVE FAILURES IN CATTLE

Heli Venhoranta

Dissertationes Scholae Doctoralis Ad Sanitatem Investigandam Universitatis Helsinkiensis

Doctoral Programme in Clinical Veterinary Medicine

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Veterinary Medicine of the University of Helsinki, for public examination in Lecture hall I, Latokartanonkaari 5, Helsinki, on 28 August 2015, at 12 noon

Helsinki 2015

Supervisors Professor Magnus Andersson
Department of Production Animal Medicine
Faculty of Veterinary Medicine
University of Helsinki
Finland

Professor Hannes Lohi
Department of Veterinary Biosciences
Faculty of Veterinary Medicine
Research Programs Unit, Molecular Neurology
Faculty of Medicine
University of Helsinki and Folkhälsan Research Center
Finland

Reviewers Professor Jørgen S Agerholm
Faculty of Health and Medical Sciences,
University of Copenhagen,
Denmark

Professor Göran Andersson
Department of Animal Breeding and Genetics
Swedish University of Agricultural Sciences,
Sweden

Opponent Professor Asko Mäki-Tanila
Department of Agricultural Sciences
University of Helsinki
Finland

ISSN 2342-3161 (print)
ISSN 2342-317X (online)

ISBN 978-951-51-1427-3 (paperback)
ISBN 978-951-51-1428-0 (PDF)
http://ethesis.helsinki.fi

Hansaprint
Vantaa 2015
Cover photo: Northern Finncattle by Reijo Jokivuori

ABSTRACT

Dairy cattle breeding programs usually rely on selection and artificial insemination (AI). The use of a small number of bulls for AI enables intense selection for desired traits but carries a risk that inherited defects, especially those controlled by recessive genes, can perpetuate in the population. Molecular genetics techniques enable efficient gene mapping even in small study cohorts. Rapid identification and management of genetic defects are crucial for preventing economic losses and maintaining good animal welfare. In this PhD study, we investigated three inherited congenital bovine defects. Two of the diseases affecting the Ayrshire breed are new and the third was described from the beginning of the 20th century in Swedish Mountain cattle.

Almost half of the pregnancies in studies I and II sired by a single Ayrshire AI bull ended in late-term abortions or stillbirth: 318 calves died. The affected calves were ~50% undersize, indicating an intrauterine growth restriction. We established that calf death resulted from a 130 kb microdeletion in the PEG3 domain on chromosome 18. The deletion truncates the 3' end of the non-coding imprinted *MIMT1* transcript and also causes expression changes in other genes of the domain. The deletion, when inherited from the sire, is semi-lethal for his progeny with an observed mortality rate of 85%. The stillbirth rate was normal when the daughters of the proband bull calved, indicating that the female mutation carriers can breed normally.

Hereditary gonadal hypoplasia in Northern Finncattle and Swedish Mountain cattle was studied in III. Our results suggest that the disease is associated with homozygosity of an ectopic segment that is duplicated and translocated from chromosome 6 to 29 (Cs_{29} allele). The same duplication is associated with colour sidedness in various cattle breeds, which coheres with the results that gonadal hypoplasia is connected with white coat colour. The duplicated segment encompasses the *KIT* gene, which is known to regulate the migration of the germ cells and precursors of melanocytes. The gonadal hypoplasia has an incomplete penetrance, which is suggested to be at

least a partially inherited trait, explaining why some unaffected animals were homozygote carriers of the Cs_{29} allele.

In study IV, we discovered an inherited disease that causes serious developmental disorder in Ayrshire cattle. The phenotype was defined as PIRM syndrome according to its typical features: **p**tosis, **i**ntellectual disability, **r**etarded growth and **m**ortality. The syndrome is autosomal, recessively inherited and caused by a G > A substitution at the last nucleotide of exon 23 in the *UBE3B* gene. Transcript analysis revealed in-frame exon skipping in the affected animals, resulting in an altered protein lacking 40 amino acids, which likely comprises protein function. Of the 129 tested Ayrshire AI bulls recently used for AI in Finland, 17% carried the mutation. Moreover, the *UBE3B* mutation may be connected with the AH1 haplotype, which is associated with reduced fertility and has a carrier frequency of 26.1% in the North American Ayrshire population. In humans, mutations in the *UBE3B* gene are associated with Kaufman oculocerebrofacial syndrome, with similar pathological effects as for PIRM syndrome.

The causative mutations of the inherited defects described are now easy to test for. The results can be used to avoid risky matings, cull carriers and provide a veterinary diagnostic. The described genotype-phenotype associations provide new insights into developmental biology and inform translational research across species. The results can also be used as a basis for candidate gene approaches to locate quantitative trait loci in cattle.

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Flisikowski K, **Venhoranta H**, Nowacka-Woszuk J, McKay SD, Flyckt A, Taponen J, Schnabel R, Schwarzenbacher H, Szczerbal I, Lohi H, Fries R, Taylor JF, Switonski M, Andersson M (2010). A novel mutation in the maternally imprinted PEG3 domain results in a loss of *MIMT1* expression and causes abortions and stillbirths in cattle (Bos taurus). PLoS One. 2010 30;5(11):e15116.

II Flisikowski K*, **Venhoranta H***, Bauersachs S, Hänninen R, Fürst RW, Saalfrank A, Ulbrich SE, Taponen J, Lohi H, Wolf E, Kind A, Andersson M, Schnieke A (2012). Truncation of *MIMT1* gene in the PEG3 domain leads to major changes in placental gene expression and stillbirth in cattle. Biol Reprod. 21;87(6):140.

III **Venhoranta H***, Pausch H*, Wysocki M, Szczerbal I, Hänninen R, Taponen J, Uimari P, Flisikowski K, Lohi H, Fries R, Switonski M, Andersson M (2013). Ectopic *KIT* copy number variation underlies impaired migration of primordial germ cells associated with gonadal hypoplasia in cattle (Bos taurus). PLoS One. 26;8(9):e75659.

IV **Venhoranta H**, Pausch H, Flisikowski K, Wurmser C, Taponen J, Rautala H, Kind A, Schnieke A, Fries R, Lohi H, Andersson M (2014). In frame exon skipping in *UBE3B* is associated with developmental disorders and increased mortality in cattle. BMC Genomics. 12;15:890.

*Equal contribution
The publications are referred to in the text by their Roman numerals.

ABBREVIATIONS

AI	Artificial Insemination
BTA	*Bos taurus*
cDNA	Complementary DNA
CN	Copy number
CNV	Copy number variant
CTS	Crooked tail syndrome
ECR	Evolutionarily conserved region
FISH	Fluorescent *in situ* hybridization
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
gDNA	Genomic DNA
GS	Genomic selection
GWAS	Genome-wide association study
IUGR	Intra-uterine growth restriction
lncRNA	Long non-coding RNA
MIMT1	MER1 repeat containing imprinted transcript 1
ncRNA	Non-coding RNA
NGS	Next generation sequencing
ORF	Open reading frame
PCR	Polymerase chain reaction
PEG3	Paternally expressed gene 3
PGC	Primordial germ cells
QPCR	Quantitative PCR
qRT-PCR	Quantitative reverse transcriptase PCR
QT	Quantitative trait
QTL	Quantitative trait loci
RT-PCR	Reverse transcriptase PCR
SNP	Single nucleotide polymorphism
UMD 3.1	University of Maryland version 3.1 of the bovine genome assembly
USP29	Ubiquitin specific peptidase 29

1 REVIEW OF THE LITERATURE

1.1 Breeding of dairy cattle and inherited diseases

Cattle keeping probably began 10,000 years ago with at least two independently domesticated cattle populations: the humpless taurine (*Bos taurus*) and the humped indicine or zebu cattle (*Bos indicus*). These interfertile cattle lines descend from the extinct wild ox or auroch (*Bos primigenius*). European taurine cattle have been subjected to intensive selection for milk and meat production and have spread almost worldwide as humans have settled new areas. Several globally important dairy cattle breeds, including Holstein, Jersey and Simmental, together with nationally important breeds such as Ayrshire and Finncattle descend from the European ancestry (1, 2).

Cattle keeping in Finland also has a long history. Cattle bones have been found dating back to 2400 BC. However, Finnish cattle breeding became important in the last half of the 19th century. One attempt to improve the cattle population was the import of foreign cattle breeds into Finland. The foreign animals were partly introgressed into the native cattle population, but pure-breeding was also valued. Nowadays the main dairy cattle breeds in Finland are Ayrshire, Holstein and Finncattle. Finncattle represent the traditional local breed, which can be divided into three sub-breeds: Eastern, Western and Northern Finncattle. The Eastern and Northern Finncattle breeds are endangered (3).

The breeding development also included systematic evaluation of animals for different traits so that the best individuals could be chosen as parents of the next generation. Several traits, such as production, breed-specific appearance and longevity were used to make breeding decisions. By the end of the 20th century the number of graded traits was multiplied and with refined statistical methods remarkably accurate breeding value estimations were produced. Using artificial insemination (AI) in breeding led to marked rates of genetic improvement and spectacular increases in productivity. In the 21st century advances in genomics have enabled genomic selection (GS),

which is replacing progeny testing. GS reduces the generation interval substantially by providing estimated breeding value at birth or even for biopsied embryos. This speeds up breeding progress to an even greater degree (2-4).

AI became a common procedure in Finnish cattle breeding in the 1950s and modern bovine breeding programs rely on it heavily (3). AI bulls with the best breeding values are used extensively and may have tens of thousands of offspring. AI enables intense selection for desired traits, but the genetic advancement of animal material may results in inbreeding. The use of a few elite animals (i.e. small effective population size) carries a risk that inherited defects, especially diseases under the control of recessive genes, can rapidly proliferate in the population. Such diseases do not become apparent until the carrier descendants of the original founder animal with the mutation are mated. Meanwhile, the deleterious mutation could have spread widely throughout the population, affected animals suddenly appearing among the population (4).

There are examples of inherited diseases that have spread to several countries with international trade in animal material. One example is Bovine Leukocyte Adhesion Deficiency (BLAD) in Holstein cattle, which is an autosomal recessive caused by a missense mutation in *ITGB2*. Calves affected with BLAD suffer from stunted growth, are susceptible to severe infections and die at a young age. However, BLAD cases have been reported in the United States, Australia, Japan, and in several countries of Europe (OMIA 000595-9913) (5, 6).

Deleterious alleles can also reach high frequencies within a few generations if the breed has a small effective population size, causing serious local problems. Mutation in the *MFN2* gene causes degenerative axonopathy in Tyrolean Grey cattle (OMIA 001106-9913) and it has been estimated that carrier frequency is close to 10% in this small breed (~5000 registered cows). Pedigree analysis revealed that one ancestor had transmitted the mutation to most of the cases and this animal was traced in 90.2% of all pedigrees (6, 7).

The spread of inherited recessives can be due to genetic drift or because of the links between causative mutations and other alleles controlling desired phenotypes. Also balancing selection maintaining a deleterious allele has been found in cattle. In such cases, the

heterozygote mutation might have beneficial effects on valuable genetic traits giving carriers a selective advantage in breeding. In Belgian Blue animals the *MRC2* mutation causes Crooked Tail Syndrome (CTS), which results in severe skeletal and muscular anomalies such as growth retardation, spastic paresis and muscular hypertrophy. The enhanced muscularity, which is a desired trait in Belgian Blue animals, was also found from carriers, explaining the high carrier frequency (~25%). Furthermore, a second mutation in the same *MRC2* gene was found with similar effects on phenotype. However, the assumed carrier frequency was much lower (~0.3%) (OMIA 001452-9913) (6, 8, 9). Recently a 660-kb homozygous deletion encompassing four genes was shown to cause embryonic lethality in Nordic Red cattle. The deletion had a dramatic effect on fertility and the carrier frequency was 13%, 23% and 32% in Danish, Swedish and Finnish Red cattle (Ayrshire), respectively. These high frequencies were accounted for by the association of the deletion with strong positive effects on milk yield and composition (OMIA 001901-9913) (6, 10).

The propagation of diseases controlled by dominant genes in cattle is not usually so widespread because the defect can be identified in the first generations. In Senepol cattle a syndrome causing lactation failure, excessively 'hairy' pelage and thermoregulatory dysfunction was caused by a dominant single nucleotide mutation in the *PRL* gene. The mutation had segregated *de novo* from a sire and its son that were used for AI (11). The Crop Ears found in Highland Cattle are caused by a dominant duplication in the *HMX1* gene. The severity of crop ears varies greatly and incomplete penetrance and/or variable expressivity of the defect have been suggested, explaining the wide spread of the dominant mutation (OMIA 000317-9913) (6, 12). Furthermore, an interesting example of a somatic mosaicism of a dominant mutation in one Charolais AI bull was found. The deletion in chromosome 2 caused Polled and Multisystemic Syndrome (PMS), which was manifested by a wide spectrum of symptoms, including death of male embryos during pregnancy (OMIA 001736-9913) (6, 13).

Even though several severe inherited bovine defects have been found the diseases can also be successfully controlled using DNA testing. Several DNA tests have been made and, for example, Semex

Alliance is testing Holstein bulls for BLAD (OMIA 000595-9913), DUMPS (OMIA 000262-9913), CVM (OMIA 001340-9913), Citrullinaemia (OMIA 000194-9913), Brachyspina (OMIA 000151-9913) and for five haplotypes related to fertility; Ayrshire bulls for Trimethylaminuria (OMIA 001360-9913) and AH1 haplotype (14). To date there are about 463 known inherited traits or disorders in cattle and for 100 of them the causal mutation has been found (6). The emergence of inherited defects is a recurrent issue in cattle breeding and management of genetic diseases requires the extension of DNA tests to allow precise identification of carriers and enable the diseases to be controlled.

1.2 Detection of causative mutations in dairy cattle

The breeding history of cattle has led to a decrease of genetic diversity and breed-specific disease heterogeneity. Unlike with human genetic diseases where different mutations cause particular syndromes, the causative mutation of a bovine defect is most likely the same in different individuals within the breed. The pedigree records that have been collected since the beginning of the 20th century are useful for estimating the mode of inheritance and relationships among studied animals. Furthermore, the large progeny of AI bulls enables an efficient comparison of affected and unaffected individuals within the family, making cattle interesting subjects in which to study inherited diseases.

Several aspects of molecular biology and bioinformatic techniques have evolved that enable efficient mutation mapping in cattle. Maybe the most important achievement was development of annotated reference sequences in 2009 (15, 16). Subsequently the maps have been upgraded. Furthermore, the progress from microsatellite markers to SNPs increased genetic marker numbers from hundreds to hundreds of thousands. The abundance and widespread distribution of SNPs in the genome, together with the technological advantages represented by SNP genotyping, have contributed to mutation mapping in cattle. In particular, the development of high density and genome-wide assays, termed SNP chips, provided an effective method for analysing the genome. The latest ground-breaking development was the emergence

of Next Generation Sequencing techniques (NGS) that enable effective sequencing of large targeted areas and even the whole genome (4, 17, 18).

The mapping of causative mutations requires several steps. The first step is usually the detection of the inherited defect and sample collection. Traditionally, study of new inherited diseases has been conducted by research teams, but collaboration with farmers, veterinarians and breeding companies is needed. In some countries the detection of inherited defects is becoming more systematic. For example, a Belgian research group led by Michel Georges has established a Heredo-surveillance platform that operates in collaboration with field veterinarians to identify emerging defects and to collect DNA samples from affected animals. Accurate phenotyping and collection of the requisite samples establishes the basis for many phenotype-genotype research projects (4, 7, 10, 12, 19, 20).

The second step in several mutation mapping projects is to trace genetic markers and haplotypes linked to the phenotype. The causative mutations are connected with the haplotype from which the mutation occurred in the founder animal. Thus the affected animals share not only the mutation but also the surrounding haplotype, and such haplotypes can be found using genome-wide linkage or association analyses. It should, however, be noted that if the mutation occurred recently, unaffected animals can be haplotype carriers even though they are free from the causative mutation (20). Case-control association analyses have been shown to be efficient for mutation mapping (7, 12) and population stratification can be avoided by using family-based association tests (11). Also a linear method can be used for association analyses (10, 21). After the association of a particular haplotype the identified loci have to be fine-mapped, usually by increasing the numbers of genetic markers in that area or by sequencing (4, 7, 11). Nowadays NGS is replacing conventional sequencing of open reading frames and exon boundaries (12, 19, 20). Furthermore, NGS of the whole genome enables research of causative mutations in a single step (22).

The candidate gene approach can also be used if the function of the gene or genes is known to be involved in the phenotype of interest. This is typically the case when a similar defect or phenotype has been

associated with a certain gene in the same or another species (9, 23). Usually, the candidate gene is studied with genomic re-sequencing (conventional or NGS), but other techniques like the gene expression analysis can be used. Furthermore, genome wide copy number variation (CNV) analyses are useful for mutation mapping and can be applied using SNP-genotyping or NGS (13, 24).

The causality of a specific mutation should be verified with additional evidence, especially when there are no earlier reports of related symptoms caused by the same gene or an orthologous gene in other species. Comparative genomics can be used for studying sequence similarity among species. Thus the conservation of an allele or particular DNA regions indicates functional importance (18). The expression analyses (RNA or protein) of the mutated genes in the affected tissue can demonstrate functional causes of the mutation (7, 19). Also cell and yeast models have been used (18, 21). A statistically significant absence of mutation in unaffected animals is also a strong proof for the causality of the mutation (18, 19), but this requires genotyping a large cohort of animals. Furthermore, supporting information about the mutation effects can be got by combining genotyping results and individual breeding data, health statistics and pedigree records (20).

Taken together, the new techniques in genetics enable effective causative mutation localization. Additionally, the broad arsenal of molecular research methods allows accurate studies of functional consequences. Results can be utilized in breeding so that screening large populations, genomic-assisted mating and culling plans can be made. Genotype-phenotype associations may also provide new insights that help biomedical research.

1.3 Imprinting and the PEG3 domain

Genomic imprinting is a rare phenomenon whereby alleles of certain mammalian genes are not functionally equivalent due to epigenetic inactivation that depends on the parental origin of the alleles. There are over 100 experimentally verified imprinted genes, based primarily on data from mice (*Mus musculus*) and humans (*Homo sapiens)* (25, 26). They represent only a small percentage of genes. According to conflict

theory, imprinting has evolved in mammals because of conflicting interests of mother and father in relation to transfer of nutrients from the mother to her offspring (27). Thus the paternally expressed genes usually increase growth, and the maternally expressed genes tend to restrict the growth of an offspring. Imprinting has an important role in mammalian placentation, development, growth and cell differentiation. Moreover, imprinted genes influence postnatal processes such as behaviour and metabolism (28-30).

Most known imprinted genes are clustered in particular chromosomal regions termed imprinted domains. Expressions of the imprinted genes in these domains are regulated by DNA methylation in the CpG rich region, the imprinting control regions (ICR). It has been shown that long non-coding RNAs (lncRNAs) are required for regulation of the imprinted expression for the whole cluster or part of it. The expression of imprinted lncRNAs is controlled through methylation (29-31).

The PEG3 domain is an imprinted gene cluster that is named after Paternally Expressed Gene 3. Other genes in this region in the cow are *ZIM2*, *AST1*, *APEG3*, *MIMT1*, *USP29* and *ZFP269* (Figure 1) (32-34). The differentially methylated region (DMR), which is assumed to be the ICR of the PEG3 domain, includes a bidirectional promoter shared by *PEG3* and *MIMT1* (or *USP29* in mouse) and part of these genes (35, 36). This DMR also includes multiple DNA-binding sites for the transcription factor YY1 in an unusual tandem repeat structure that covers the first intron of *PEG3* (37, 38).

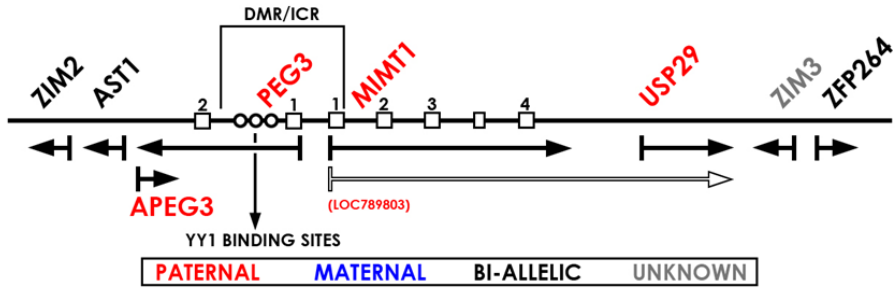

Figure 1. The illustrated structure of the bovine PEG3 domain which includes at least eight genes. The differentially methylated region (DMR) supposed to be the imprinting control region (ICR) of the domain includes the bidirectional promoter shared by *PEG3* and *MIMT1* and part of these genes. The first intron of *PEG3* includes several binding sites for the transcription factor YY1.

In the middle of the PEG3 domain is a 250 kb region lacking any obvious open reading frame (ORF), but including several evolutionarily conserved regions (ECRs) (34, 39-41). These ECRs are likely *cis* regulatory elements that may be involved in controlling the transcription and imprinting of the PEG3 domain (39). It has been hypothesized that transcription factor YY1 might link the ECRs and ICR of the PEG3 domain, enabling long-distance interaction between the ECRs and the bidirectional promoter of *PEG3* and *MIMT1/USP29* (42).

The PEG3 domain is well conserved among different mammal species but the protein-coding capacity of several genes has been lost during recent evolution. According to existing results, the *PEG3* gene is the only gene that has maintained its protein-coding capacity in all lineages of mammals. This indicates that functioning of PEG3 protein is essential. Nevertheless, the RNA genes of the PEG3 domain are still transcriptionally active, indicating that these genes might have functionally adapted as non-coding RNA (ncRNA) genes with possible regulatory functions (42).

MIMT1 (also called *ITUP1* or *IMPO1* in humans) is one of the ncRNA genes of the PEG3 domain in humans and cows. The *MIMT1* is localized in the well conserved middle region of the PEG3 domain, between *PEG3* and *UPS29* genes (36). Very little is known about *MIMT1*. Earlier results showed that it has five exons and four

alternatively spliced transcripts in cattle (34, 41). However, it has been proposed that, as in mouse, the bovine *MIMT1* and *USP29* might share exons (42). Furthermore, the predicted *LOC789803* gene in the NCBI database covers the whole middle region of the PEG3 domain. *LOC789803* has nine predicted transcripts of which several are aligned to *MIMT1* mRNA sequences found in earlier studies (34, 41). Even though the function of *MIMT1* remains unknown, the location of the promoter in the ICR and conservation of the whole gene area indicate that *MIMT1* may have a vital role in mammalian survival.

1.4 Gonadal hypoplasia in Swedish Mountain cattle and white coat colour

Gonadal hypoplasia appears as aberrant small size and underdevelopment of ovaries and testicles, which leads to fertility problems, especially if both gonads are affected. Examples of different types of gonadal hypoplasia in several mammalian species have been reported (43-50) and perhaps the most studied is the inherited gonadal hypoplasia in Swedish Mountain cattle (also referred to as the Swedish Highland breed). The defect is old; it emerged already in the 20th century simultaneously with pure breeding of Swedish Mountain cattle (51, 52). Later the incidence of gonadal hypoplasia increased substantially but systematic clinical investigations and removal of affected animals from breeding led to the successful reduction of the defect. The prevalence decreased from 17.3% to 7.3% in seventeen years (53). Gonadal hypoplasia is also found in Northern Finncattle. The frequent use of Swedish Mountain cattle for breeding Northern Finncattle indicates that the defect might have been introduced to Finland from Sweden.

The gonadal hypoplasia of Northern Finncattle and Swedish Mountain cattle is a congenital defect (52) and study of foetal ovaries revealed that the hypoplastic ovaries can already be identified at the foetal stage, which could indicate a failure of the migration and synchronous mitotic divisions of primordial germ cells (PGC) (51). The comprehensive breeding experiments in Swedish Mountain cattle indicated an autosomal recessive mode of inheritance with incomplete penetrance (0.5). Furthermore, it has been suggested that the

incomplete penetrance is partly an inherited feature (52). The proportions of left-, double- and right-sided gonadal hypoplasia are 82%, 15% and 3%, respectively, showing that the defect is manifested mainly in the left side (52). The predominance of the left side was confirmed in other studies (54, 55), but there is no evidence that side of the defect could be genetically determined (52). Also the severity of gonadal hypoplasia is known to vary from total (no germ cells) to partial (reduced number of germ cells) (51, 52). Animals with bilateral total gonadal hypoplasia are sterile. Moreover, the secondary sexual characteristics of bilaterally affected animals can be changed in females because of the impaired production of sexual hormones (51, 52). In males similar changes have not been noticed (52), most probably because Leydig cells can produce testosterone also in affected testicles. No other health problems have been reported for animals that suffer from gonadal hypoplasia.

Gonadal hypoplasia has been associated with white coat colour in Northern Finncattle and Swedish Mountain cattle (51). The coat colour of the breeds varies from total white to total black or brown with numerous intermediates. The most common colour pattern of Northern Finncattle is white coat with pigmented ears and muzzle, possibly together with spotted sides and coloured legs (Study III/Figure S1). None of the affected animals were over 40% coloured (Study III/Table S1) (51). The colour variation is partly due to the colour-sided pattern (56, 57) that is determined by two loci present on BTA29 and BTA6. The Cs_{29} allele on BTA29 resulted from duplication and translocation of a 492 kb segment of BTA6 including the *KIT* gene. The Cs_{6} allele, residing on BTA6, is a result of a subsequent duplication and translocation that moved the segment comprising fused sequences of the BTA29 and BTA6 back to the *KIT* locus in BTA6 (Figure 2). It is indicated that in both cases the dysregulation of the *KIT* gene leads to the colour sidedness (24). The *Cs* alleles have been associated with colour-sided patterns in several bovine breeds, including White Galloway, White Park, Belgian Blue and the yak (23, 24, 58).

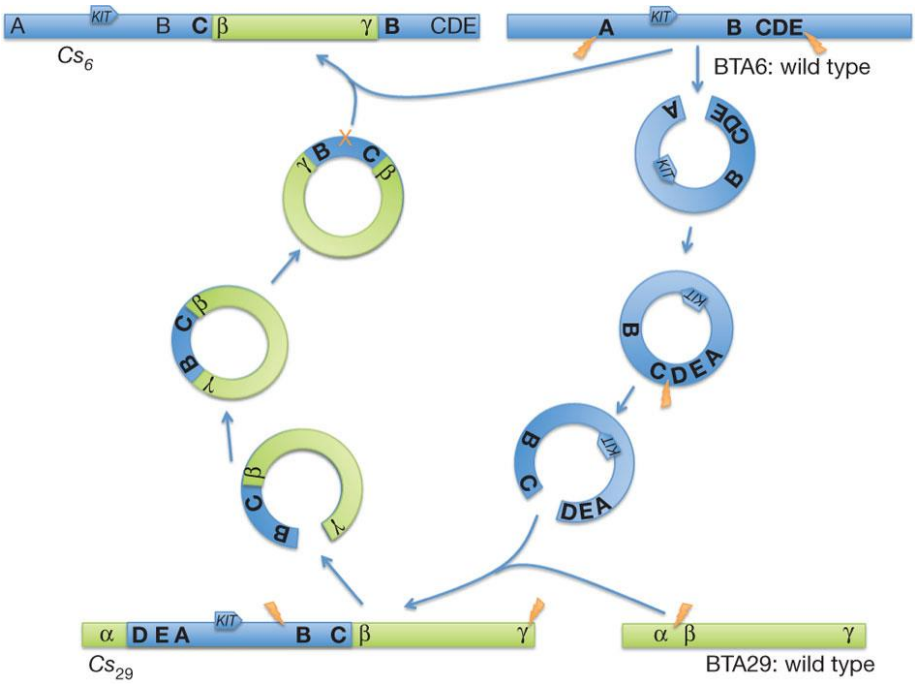

Figure 2. Model for the generation of the Cs29 and Cs6 alleles by serial translocation via circular shuttling intermediates. Modified and used with permission (24).

In cattle also the spotted haplotype (piebaldism) has been associated with BTA6, indicating the *KIT* gene to be a candidate for spotting locus (59-62), especially with the white colour of the face in the Hereford and yak (23, 60). *KIT* mutations also underlie coat colour variation in other mammals, e.g. pig (*Sus scrofa)* (OMIA 000209-9825, 001743-9825 and 001216-9825), cat (*Felis catus)* (OMIA 000214-9685, 000209-9685) and horse (*Equus caballus)* (OMIA 000209-9796) (6). Furthermore, in cats the white colour-causing mutation in *KIT* is associated with deafness and iris hypopigmentation, which can be uni- or bilateral with incomplete penetrance (63, 64). In humans, *KIT* mutations cause piebaldism, mast cell disease and several types of tumour (*164920) (65). In mice, often pleiotropic mutations in *KIT* result in impaired pigmentation, reduced fertility or sterility, anaemia and deafness (MGI:96677) (66).

Several studies have shown that normal function of KIT protein is crucial for the survival, proliferation and migration of the PGCs and

melanoblasts (melanocyte precursors) (67-70). These cell types develop from the pluripotent neural crest of the embryo and migrate along characteristic pathways to their destination tissues. The germ cells colonize the gonads and melanocytes reside in the skin, hair follicles, inner ear, and parts of the eye (70-73). Furthermore, *KIT* is also essential for the generation of hematopoietic cells (67, 74). The association of *KIT* with the development of several cell types explains various symptoms and pleiotropic effects caused by *KIT* mutations.

1.5 PIRM syndrome in Ayrshire cattle and ubiquitination

An increasing number of Ayrshire calves with severe developmental defects that often lead to death were identified in Finland during 2011–2014. The clinical examinations of affected animals combined with lineage studies indicated an inherited disease with a complex phenotype. The disorder was classified as PIRM syndrome after its prevalent symptoms: **p**tosis, **i**ntellectual disability, **r**etarded growth and **m**ortality.

The bovine PIRM syndrome resembles Kaufman oculocerebrofacial syndrome (KOS), also called blepharophimosis-ptosis-intellectual disability syndrome in human (MIM 615057, MIM 244450) (65). The patients present severe developmental delay combined with ocular and other craniofacial anomalies and multi-organ abnormalities. Patients with KOS have biallelic mutations in the *UBE3B* gene, which encodes an ubiquitin E3 ligase protein (75-78). Several other malfunctions of E3 ligases are associated with a variety of human developmental diseases (79-82). The best known example of E3 ligase related defects is Angelman syndrome, characterised by intellectual disability, absence of speech, motor dysfunction and seizures (MIM 105830), which are caused by mutations on the imprinted gene *UBE3A* (83, 84).

The E3 ligase proteins are a large protein family that play key roles in the recognition of protein substrates for ubiquitination. The protein ubiquitination is a post-translational protein regulation pathway related to protein degradation and several other molecular processes essential to normal neurodevelopment and organogenesis. Besides the E3 ligase,

the transfer and covalent attachment of ubiquitin to target proteins requires the activating enzyme (E1) and a conjugating enzyme (E2) (Figure 3). The UBE3B protein belongs to HECT (**h**omology to **E**6-AP **C**-**t**erminus) domain-containing E3 ligases, which besides the C-terminal HECT have a variable N-terminal region conferring the ability to bind specifically to their substrates. HECT domains consist of two subdomains, the N-terminal subdomain, which contains the E2 binding site, and the C-terminal subdomain that harbours the catalytic Cys residue required for ubiquitin transfer to the substrate. Thus the HECT domain is the active site of this type of E3 ligase (85, 86). Consistently almost all KOS patients have mutations that likely compromise or eliminate the catalytic activity of the HECT domain, indicating a strong interference of UEB3B ligase function (75-78).

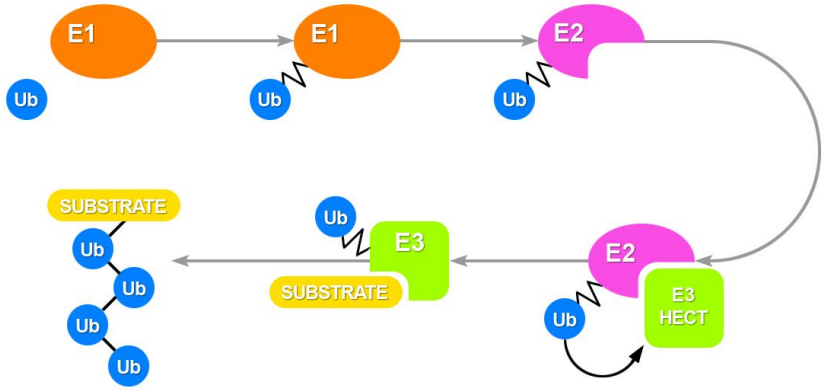

Figure 3. The ubiquitin pathway including ubiquitin-activating (E1), ubiquitin-conjugating (E2) and HECT-type ubiquitin-protein ligase (E3) enzymes.

2 AIMS OF THE STUDY

This study targets three congenital conditions related to reproduction in cattle based on hypotheses that they are all hereditary due to their preponderance in particular breeding lines. The aim was to define two new inherited diseases in Ayrshire cattle and to identify their genetic causes and those of a defect already found in Northern Finncattle and Swedish Mountain cattle. A further objective was to explain the cause of the mutations at the RNA level, including broader transcription studies.

The specific aims of the projects were:

I Define a new disease causing a high rate of late abortions and stillbirths in the offspring of one Ayrshire AI bull and study the genetic etiology of the defect.

II Determine the borders of the deletion causing late abortions and stillbirths in Ayrshire cattle and analyse the expression and methylation changes in the placenta.

III Identify the genetic cause of the inherited predominantly left-sided gonadal hypoplasia in Northern Finncattle and Swedish Mountain cattle.

IV Describe the symptoms of the new syndrome causing developmental defects and mortality in Ayrshire calves and identify the causal mutation.

3 MATERIALS AND METHODS

3.1 Ethics statement (I-IV)

Blood sampling, clinical studies and insemination were carried out using standard Finnish veterinary protocols. All animal experiments were approved by the Animal Ethics Committee of the State Provincial Office of Southern Finland (STH051A, ESAVI-2010-08583/YM-23, ESAVI-2010-03428/Ym-23, ESAVI/641/04.10.07/ 2014).

3.2 Clinical examination and sampling (I-IV)

In study I semen analysis of the proband bull included measurement of sperm concentration, total sperm content and motility studies of each ejaculate and post-thaw motility studies. A smear test was made from two ejaculates for morphological examination of spermatozoa. The early fertility data were evaluated using the estimated breeding values for non-return rate (within 60 days of each insemination) for the inseminations. Heparin blood of the proband bull was collected for cytogenetic analyses. Seven dead calves were subjected to necropsy in study I.

In studies I and II semen of the affected bull was used for the insemination of cows scheduled for slaughter. Samples from cotyledon (placental structure of the foetal side), caruncle (placental structure of the maternal side) brain, lung, kidney, heart, liver, and muscle were collected for DNA and RNA studies. Histological samples were taken from the cotyledon and caruncle, which were subjected to standard formalin fixation and embedded in paraffin. Sections (5 μm) were cut and stained with haematoxylin-eosin. Samples from 22 foetuses were collected between 41–157 days of pregnancy. In total, 122 DNA samples were analysed in studies I and II.

In study III clinical examinations of gonads were done during farm visits by experienced veterinarians. Males were palpated for testicle size and symmetry and females older than 16 months, excluding animals more than five months pregnant, were studied using ovarian

palpation per rectum. In post-mortem studies, gonads were examined visually, palpated and weighed. Also histological samples were collected and embedded in paraffin after standard Bouin's fixation. These samples were cut and stained as earlier. DNA samples from 96 animals were included in study III. Also heparin blood for cytogenetic analyses was collected from five animals.

In study IV most of the clinical examinations and symptom observations were done in farms by local veterinarians, farmers and breeding advisers. Study IV included DNA samples from 188 animals. RNA samples were collected post-mortem from five animals' cerebral cortex, tectum, hippocampus, cerebellum, lung, liver, heart, kidney, spleen and ovary tissues.

3.3 DNA and RNA isolation (I-IV)

For study I DNA was extracted with a DNeasy Blood and Tissue kit (Qiagen) according to manufacturer's instructions. In studies II-IV a semi-automated Chemagen extraction robot (Chemagen Biopolymer-Technologie AG) was used for DNA isolation from blood samples. Standard protocols for proteinase K digestion and phenol-chloroform extraction were also used for blood sample DNA extractions in study IV. DNA samples from hair bulbs were lysed as previously described (87) in study I.

DNA from frozen and diluted semen samples was extracted with a Qiagen Kit (QIAamp DNA Mini Kit) and a Chemagen extraction robot (Chemagen Biopolymer-Technologie AG) in studies III and IV respectively. Extraction with a Qiagen Kit was made according to the DNA Purification from Tissues-protocol in QIAamp DNA Mini Kit handbook with some modifications. In brief, 200–500 µl of semen was washed with 200 µl of phosphate buffered saline (PBS) (centrifugation for 5 min at 10,000 g) and the Qiagen buffer ALT was added up to 300 µl together with 20 µl of Proteinase K and dithiothreitol. During incubation (1 h at 56 °C) the sample was pulse vortexed four times for 15 sec. After adding 300 µl of Qiagen buffer AL the sample was incubated for 10 min at 56 °C and thereafter 150 µl of 96% alcohol was added. The sample was incubated for 3 min at room temperature. The whole mixture was applied to the QIAamp Mini spin column and

centrifuged at 6,000 g for 1 min, washed twice with 500 µl of Qiagen buffer AW1 and once with 500 µl of Qiagen buffer AW2 (centrifugation at 6,000 g for 1 min). All filtrates were discarded. Before elution, the column was dried with centrifugation at 20,000 g for 3 min. Lastly 50 µl of distilled water was added into the column, incubation was for 1 min at room temperature and centrifugation at 20,000 g for 1 min.

Semen extraction with a Chemagen extraction robot was also started with a wash: 200 µl of semen was washed twice with 1000 µl PBS (centrifugation for 5 min at 10,000 g). The pellet was resuspended in 500 µl lysis buffer (Chemagic DNA Blood Kit special, article No. CMG-703-1) containing 2 µl proteinase K (20 mg/ml) and 20 µl DTT (1 M). After overnight incubation at 55 °C extraction was continued according to the manufacturer's instructions with 1 ml isolation buffer and 150 µl elution volume. Proteinase K digestion and phenol-chloroform extraction were also used for DNA isolation from semen in study IV.

DNA was extracted from tissue samples with a DNeasy Blood and Tissue kit (Qiagen) in study I and with a QIAamp DNA Mini Kit (Qiagen) in studies II and III. Total RNA in studies I and II was extracted using Trizol (Invitrogen) and in studies II and IV with an RNeasy Mini Kit (Qiagen). Extractions were made according to manufacturers' instructions from tissue samples. RNA was converted to cDNA with First Strand cDNA Synthesis Kit (Fermentas) in studies I and II. A high Capacity RNA-to-cDNA Kit (Applied Biosystems) was used in study IV.

3.4 Cytogenetic analysis (I and III)

Lymphocyte cultures were established from the carrier sire in study I and from five Finncattle in study III. In study I, slides with fixed lymphocytes were Giemsa-stained and G-banded and the international nomenclature for bovine chromosomes was followed (88). FISH analysis in study III was carried out as described by Durkin *et al.* (24). In brief, BAC clones RP42-160M9, RP42-156I13, RP42-37P11 and RP42-116G8 were derived from the RPCI-42 Bovine BAC Library (89). After alkaline lysis extraction BAC DNA was labelled by random

priming and clones RP42-160M9; RP42-156I13 and RP42-37P11, RP42-116G8 were mixed and labelled again with biotin-11-dUTP and digoxigenin-11-dUTP respectively. These probes were separately denatured (10 min at 70°C) with an excess of bovine Cot-1 DNA and hybridized with denatured chromosome slides (overnight at 37°C). Biotin-labelled probes were detected using streptavidin-Cy3 (Amersham, 1:200, red colour) and digoxigenin-labelled probes were detected with anti-digoxigenin-fluorescein Fab fragments (Roche, 1:200, green colour) from washed slides. Counterstaining was done with Vectashield containing DAPI (Vector Laboratories) and slides were examined with an epifluorescence Nikon E600 Eclipse microscope equipped with a cooled digital CCD camera and Lucia software.

Synaptonemal complex analysis in study I was performed as earlier described (90, 91) immediately after slaughter of the proband bull.

3.5 SNP genotyping

3.5.1 Genotyping and quality control (I, III AND IV)

Arrays used for genotyping were BovineSNP50 BeadChip (Illumina) in study I and BovineHD BeadChip (Illumina) in studies III and IV. The default parameters of Illumina's BeadStudio were used for genotype calling and chromosomal positions were determined on the basis of the University of Maryland reference sequence UMD3.1 (92). In studies III and IV Y-chromosomal (1224 SNPs), mitochondrial (343 SNPs) and SNPs with an unknown chromosomal position (1735) were excluded from further analysis. The quality control was carried out with PLINK (93). The genotypes of two and one animals were omitted because genotyping failed in more than 10% of the SNPs in studies III and IV respectively. Further, 6229 SNPs in study III and 7235 SNPs in study IV were excluded because genotyping failed in more 10% of the individuals. Moreover, 121,657 SNPs in study III and 149,129 SNPs in study IV were deleted because SNPs were monomorphic.

3.5.2 Genome-wide association study (I, III and IV)

GridQTL (94) was used for genome-wide linkage analysis and custom software was used for the allele frequency association model in study I. Haplotypes for the associated region on BTA18 were constructed manually and positions of the SNPs were based on the UMD3.1 assembly (92) in study I.

Genome-wide allelic and genotypic associations in study III were analysed with Fisher exact tests by using PLINK (93) and SNPs with $P < 7.71 \times 10^{-8}$ were considered as significantly associated (Bonferroni-corrected threshold for multiple testing). The extent of false positive association signals was assessed as earlier (95) by inspecting quantile-quantile plots and calculating genomic inflation factors in study III.

Imputation of sporadically missing genotypes and haplotype inference analysis was done with Beagle genetic analysis software (version 3.2.1) (96) in study IV. Steps of 15 SNPs were used for shifting of 80 adjacent SNPs containing a sliding window along the entire genome. The allelic association of haplotypes with a frequency > 5% within each window were analysed with Fisher's exact tests.

3.5.3 Detection of copy number variants (III)

In study III copy number variations were analysed from SNP genotype signal intensities after quality control with PennCNV (97). In brief: CNV-detection algorithm considers several variables (log R ratio, B allele frequency, allele frequency and distance of adjacent SNPs) and Fisher exact tests were used for the comparison of CNVs (minimum number of 10 SNPs and minimum length of approximately 35 kb) between control and cases.

3.6 Fine mapping and validation of mutations

3.6.1 Next generation sequencing (IV)

TruSeq DNA sample preparation kit (Illumina) was used for the preparation of the libraries and sequencing was done with the HiSeq 2000 system (Illumina, San Diego, CA, USA). Reads (length 101 bp) were processed during the sequencing step with the Illumina BaseCaller. Burrows Wheeler Aligner (version 0.6.1-r104) (98) with its default parameters was used for alignment of the reads to UMD3.1 assembly (92). SAM files (Sequence Alignment/Map) were converted into BAM files (Binary Alignment/Map) with SAMtools (version 0.1.18) (99) and duplicate reads were identified and marked with the MarkDuplicates command of Picard (100).

The multi-sample approach implemented in mpileup of SAMtools with BCFtools (99) was used for polymorphism calling the region of Chr17:60,000,000 bp – 70,000,000 bp. Duplicated reads and positions with coverage over 720 reads (corresponding to 2 x N samples x average coverage) were excluded from the variant calling. The Beagle (version 3.2.1) phasing and imputation was used to improve the primary genotype calling by SAMtools.

Variants found with the multi-sample variant calling in the extended homozygosity segment (Chr17:65,645,831 bp – 66,358,629 bp) were analysed according to the presumed recessive mode of inheritance. UMD3.1 bovine genome assembly (101) was used for the prediction of the candidate causal variants' functional effects.

3.6.2 PCR and Sanger sequencing (I - IV)

Primers (Study I/Table S4, Study II/Table S1, Study III/Chapter "PCR and sequencing" and Study IV/Table S4) were designed using Primer3 software (102). In studies I, III and IV PCR reactions contained 20 ng of genomic DNA (gDNA), 1.5 mM of $MgCl_2$, 1× PCR buffer (Qiagen), 0.5 µM of both primers, 200 µM of each nucleotide, and 0.5 units of Taq DNA polymerase (Qiagen). In study II PCR reactions contained 100 ng of genomic DNA, 3 mM of $MgCl_2$, 1× Green GoTaq

reaction buffer (Promega), 0.5 µM of both primers, 250 µM of each nucleotide and 5 units of GoTaq polymerase (Promega). Thermal cycling conditions were: initial denaturation at 95 °C for 3–5 min, followed by 35 cycles of denaturation at 95 °C for 30–40 sec, annealing at 58°–64 °C for 40–60 sec, elongation at 72 °C for 1 min and then a final elongation at 72 °C for 2–3 min. Sizes of the PCR products were confirmed with 1.5 % agarose gel electrophoresis. Purified PCR products were Sanger sequenced with a 3730xI DNA Analyzer (Applied Biosystems). Phred (103-105), Consed (106), Sequence Scanner 1.0 software (Applied Biosystems), T-Coffee (107) and Variant Reporter 1.0 (Applied Biosystems) software were used for sequence processing and analysing the results.

3.6.3 Quantitative PCR (I and III)

Primer sequences designed using the Primer3 program are given in Study I/Table S4, and Study III/Chapter "QPCR". In study I semi-quantitative copy number detection was done with standard PCR and primer pairs in deletion and control areas. Amplicons were loaded on the 1.5% agarose gel and the band intensities were visualized using GelDoc System (Intas).

In study I qPCR was performed with two different primer pairs in the deletion area and two control primer pairs. The relative copy number for each target region was calculated as $2^{\wedge(1+\ (-\Delta\Delta ct))}$. The corrected CTs from all tested samples tended to form two discrete clusters.

In study III two primer pairs in CNV area of BTA6 and one control primer pair were used for qPCR. A method based on Weksber *et al*. (108) and Lachman *et al*. (109) was applied for analysis.

All qPCR reactions were done in triplicate. SYBR® Green detection chemistry (Life Technologies) and Applied Biosystems® 7500 Real-Time PCR systems (Life Technologies) were used for the qPCRs according to manufacturer's recommendations.

3.7 Frequency analysis of the mutations (II, III and IV)

In studies II and III simple PCR tests for mutation screening were used. PCR conditions were described earlier. In study II amplicons were produced with primer pairs that spanned the deletion area and with control primer pairs outside the deletion area (Study II/Table S1). In each assay, four positive control samples and four negative control samples were used. In study III breakpoint primers designed by Durkin *et al.* (24) together with one primer pair that flanked the insertion site of the wild type BTA29 (Study III/Chapter "PCR and sequencing") were used for CNV testing. Product sizes were tested in 2% and 1.5% agarose gel electrophoresis in studies II and III respectively.

In study IV KASP (Kompetitive Allele Specific PCR) reagents (LGC) and a 7500 Fast Real-Time PCR instrument (Applied Biosystems) were used for mutation frequency analysis according to the manufacturer's instructions. The qualities of every run were verified with two samples of each polymorphism group that had been tested earlier by Sanger sequencing.

3.8 Luminometric methylation assay (II)

In study II global DNA methylation was analysed with a luminometric methylation assay as previously described (110). In brief, gDNA (1 µg) was cleaved with restriction enzymes *Hpa*II and *Eco*RI or with *Msp*I and *Eco*RI (FastDigest; Fermentas) in two separate reaction mixtures in a 24-well format using a PyroMark Q24 system (Qiagen). Luminometric peak calculation was done with PyroMark Q24 software. The ratios of both restriction enzyme pairs were calculated as (dGTP + dCTP)/aATP for their respective reactions. DNA methylation was calculated from the *Hpa*II:*Msp*I ratio, where a ratio of 1 indicates 0% methylation and a ratio approaching 0 corresponds to 100% DNA methylation.

3.9 Manual re-annotation of the bovine *UBE3B*-gene (IV)

In study IV, the genomic structure of *UBE3B* was re-annotated with the *GENOMETHREADER* software tool (111) based on the UMD3.1 bovine genome sequence assembly (92) and the Dana-Farber Cancer Institute bovine gene index release 12.0 (112). The resulting output was viewed and edited using the Apollo sequence annotation editor (113). The effect of the p.E692E-polymorphism on mRNA splicing was predicted using the web-based tool ESEfinder 3.0 (114).

3.10 RNA Expression analysis

3.10.1 Reverse transcriptase PCR (I, II and IV)

In studies I, II and IV gene expression was analysed from several tissues with reverse transcriptase PCR (RT-PCR), with similar reactions as earlier. In study I the expression of *MIMT1* was analysed with one primer pair (Study I/Table S4) and in study II the *NPSR1* expression was studied with two primer pairs of which one was specific for *NPSR1* isoform A (Study II/Table S1). *GAPDH* was used as an endogenous control in both studies. *UBE3B* gene expression was analysed with two primer pairs (Study IV/Table S1) in study IV. The intensities and sizes of PCR products were compared with 1.5% or 2% agarose gels and identities of the products were confirmed by sequence analysis as earlier. The study IV PCR products obtained with primer pair 2 were extracted from the gel with a GenEluteTM Gel Extraction Kit (Sigma-Aldrich) before sequencing.

3.10.2 Microarray (II)

In study II, the RNA expression of several thousand genes was studied with the SurePrint G3 custom gene expression microarray platform (8×60 k, AMADID 031042; Agilent). The array was designed based on the Gene Expression Microarray (AMADID 023647; Agilent) and

additional transcripts from Entrez Gene (October 2010) and Ensembl59. Low-Input Quick Amp labelling kit, one-color (Agilent) was used for Cy3-labeled cRNA production. The cDNAs were hybridized to the microarray slides according to the manufacturer's instructions and washed. Scanning was done with a DNA microarray scanner (G2505C model; Agilent) at 2-μm resolution and image processing was performed with Feature Extraction software version 10.7.3.1 (Agilent). Filtration of processed signals was based on ''Well above background'' flags, that is, detection in three of four or four of six samples in either of two experimental groups, and normalized with BioConductor software VSN (variance-stabilizing normalization) (115). Normalized data were analysed with a distance matrix and a heat map based on the pairwise correlation of the samples (gene plotter software; BioConductor) for quality control. Limma software (BioConductor) (116) was used for significance analysis and method of false discovery rate was used for the correction of multiple testing.

3.10.3 Quantitative real-time reverse transcriptase PCR (I and II)

In studies I and II qPCR experiments of gene expression were performed using the same chemistry and equipment as earlier. Melting curve analysis was used for primer specificity (Study I/Table S3 and Study II/Table S1) and capture temperatures determination. The needed threshold cycle (CT) number was calculated using the second derivative maximum method. The CT is correlated inversely with the logarithm of the initial template concentration. The relative expression difference between the analysis groups was calculated for each animal (ΔΔCT). All cDNA samples were assayed in triplicate, and relative expression levels were normalized to endogenous *GAPDH* expression.

3.11 Protein analyses (II and IV)

3.11.1 Western blot (II)

Frozen foetal cotyledon samples were lysed using mechanical homogenization in Mammalian Cell Lysis buffer (Sigma). Quantification of protein samples was done using Advanced Protein Assay reagent (Cytoskeleton). Proteins were separated by 12% SDS-PAGE with a 10–40 μg of total protein in each lane and electroblotted onto polyvinylidene fluoride membranes (Immobilon-P; Millipore). After blocking with 5% milk, membranes were incubated overnight at 4°C with the primary antibody anti-NPSR1 (1:500 dilution; Sigma) and anti-GAPDH (1:5000 dilution; Sigma). Incubation with the horseradish-peroxidase-labelled secondary antibody anti-rabbit (1:5000 dilution; Sigma) and anti-mouse immunoglobulin G (IgG; 1:6000 dilution; Abcam) was done for 1 h at room temperature. Chemiluminescence was detected with ECL Western blotting substrate kit (Pierce) using x-ray film (Agfa).

3.11.2 Simulation of the protein (IV)

The ClusterW2 tool (117) was used for protein alignment and the effect of the absence of exon 23 on protein structure was investigated using the protein homology recognition engine V2.0 - PHYRE2 (118).

4 RESULTS

4.1 A deletion in the imprinted PEG3 domain results in a loss of *MIMT1* expression and causes late-term abortions and stillbirths in cattle (I and II).

4.1.1 Half of the pregnancies sired by the proband bull ended in stillbirths or abortions (I)

The semen of a Finnish Ayrshire proband bull was commercially used for AI of 1,900 heifers and cows in 2006 and 2007. One year later various farmers began to report late gestation abortions and stillbirths of the pregnancies sired by the bull. In total 318 calves died (42.6% of all offspring). The corresponding average percentage of late abortions and stillbirths for the AI bulls of the Ayrshire breed in Finland is 5%. Dead calves were ~50% undersized, indicating intrauterine growth restriction, and had uninflated lungs, but otherwise they were phenotypically normal according to necropsy. Ten crossbreed pregnancies with Holstein females resulted in three live-born calves and seven either stillborn or aborted calves after at least seven months of gestation. The stillbirth rate was normal (4%, 133 calvings) when the daughters of the proband calved. No male offspring were used for breeding.

The semen and fertility parameters of the proband bull were within normal ranges and the bull had a normal karyotype (60, XY) with no abnormality detected. No chromosomal rearrangement in any of the studied primary spermatocytes was found with synaptonemal complex analyses.

The results indicated that a phenotypically normal bull transmitted a lethal allele to approximately 50% of its offspring and unaffected daughters of the proband bull could breed normally. Furthermore, no chromosomal structural rearrangement was found. It was hypothesized

that the causative mutation localises in a maternally imprinted gene that is expressed only when inherited from the sire.

4.1.2 Late-term abortions and stillbirths are associated with the distal end of chromosome 18 (I)

To identify the genomic region associated with stillbirths and abortions, 42 tissue samples were SNP genotyped with BovineSNP50 BeadChip genotyping array (Illumina). This included also eight foetuses of which dams were inseminated with the semen of the proband bull before slaughter. The linkage analysis was done according to the assumption that the causative mutation was in a maternally imprinted gene, indicating that linkage disequilibrium information concerning the location of the disease locus could be extracted only from the paternally inherited alleles. The genome-wide half-sibling linkage analysis was done with five affected calves and 13 unaffected calves, all offspring of the proband sire. The linkage analysis included 15,631 autosomal loci in which the sire was heterozygous. The analysis revealed the association locus at the BTA18 ($P = 0.0618$, Figure 4A). Linkage analysis and allele frequency association analysis performed for individual SNPs in BTA18 localized the mutation to the distal end of the chromosome 18 (Figure 4B).

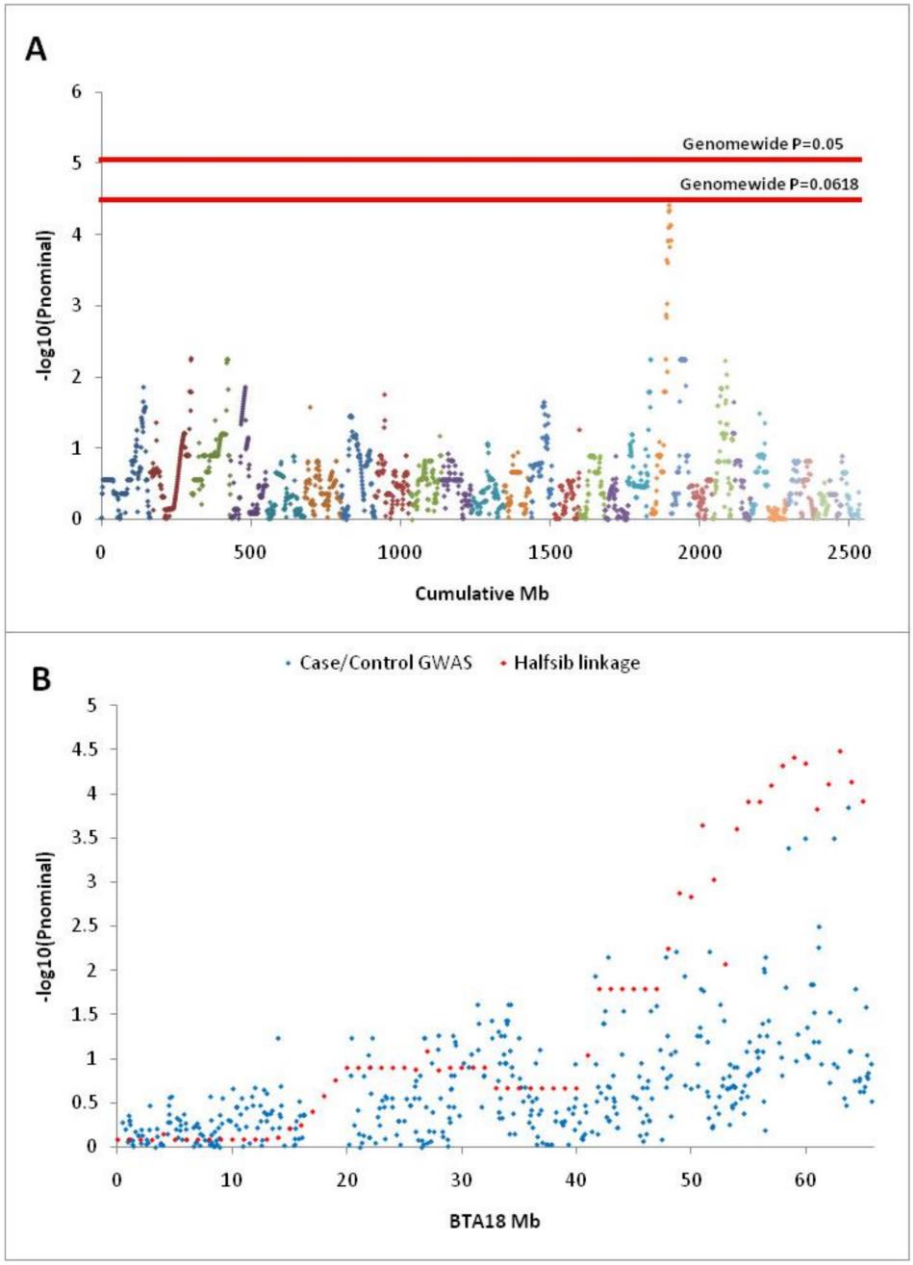

Figure 4. **(**A) Genome-wide half-sibling linkage analysis localizes the casual mutation to BTA18 at a genome-wide significance level of P = 0.0618. Plot resolution is 1 cM (assumed equivalent to 1 Mb) and phenotypes of 1 were assigned to each of five affected calves and 0 to each of 13 unaffected calves. (B) Linkage analysis and allele frequency association analysis performed for individual BTA18 SNPs localizes the mutation to the distal end of the chromosome.

The SNP haplotype analysis of the identified area was done with six affected calves, 13 unaffected calves, eight foetuses together with the sire, grandsire and 13 dams. The haplotype analysis included only the SNPs in which the proband sire was heterozygous. All affected calves had inherited the same haplotype region from the proband bull (Study I/Table S1 and Figure 3). The same haplotype was present in two unaffected calves and four foetuses. During the haplotype construction, we found Mendelian misinheritance between proband sire and his affected offspring at two SNPs. The genotypes of ss86313444 and ss86322441 SNPs were confirmed with sequencing SNP loci and the sequence surrounding them (Table 1). These SNPs were hypothesised to be heterozygous null alleles in the proband bull and some of his progeny, indicating presence of deletion. The assumed deletion region harbours the imprinted PEG3 domain.

Table 1. Genotypes of the null SNPs were re-analysed with Sanger sequencing. Offspring showing Mendelian misinheritance are highlighted in light grey, indicating that these individuals inherited the deletion allele from the proband sire (YN51). A and B in sample code indicate affected and unaffected half-siblings respectively. C indicates foetuses sired by YN51.

Sample	ss86313444	ss86322441
YN51	**G-G**	**C-C**
A2	A-A	C-C
A3	A-A	T-T
A4	A-A	T-T
B13	A-G	not studied
B14	A-G	not studied
C1	A-A	T-T
C2	A-G	C-C
C3	A-A	T-T
C4	A-G	C-T
C5	G-G	C-C
C6	G-G	C-C
C8	A-G	C-T

4.1.3 Microdeletion removes part of the imprinted PEG3 domain (I and II)

To confirm the presence of a deletion and to evaluate its size we constructed an additional dense panel of 79 SNPs in and around the region of interest (Figure 5, Study I/Table S2) applying a comparative sequencing approach. Eight foetuses and the proband bull were studied. The misinheritances and loss of heterozygosity (LOH) at 20 of the SNPs in carrier foetuses and proband bull confirmed a heterozygous microdeletion approximately spanning 64325122 and 64431506 bp. Due to the distance between SNP markers the estimated size of the deletion can range between 110 and 130 kb in the analysed animals.

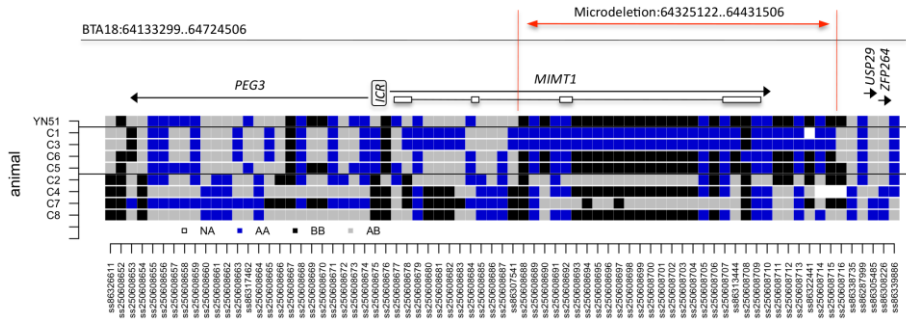

Figure 5. Foetuses C1-C8 are offspring of the proband sire YN51. Foetuses C1, C3, C5 and C6 have inherited the lethal deletion and foetuses C2, C4, C7 and C8 have inherited the normal allele. A red line denotes the microdeletion region. ICR indicates the imprinting control region. *MIMT1* is structured into 4 non-coding exons (open boxes). The genomic annotation of the region is based on the UMD3.1 assembly (92). SNP information is given in Study I/Table S2.

The presence of the deletion was also analysed with a quantitative PCR-based copy number detection approach in eight foetuses and the proband bull. The comparison of signal intensity between putative carrier group (sire, C1, C3, C5 and C6) and putative non-carrier group (C2, C4, C7 and C8) confirmed the presence of the microdeletion (Study I/Figure 6).

Finally, a series of direct sequencings was used to identify genomic breakpoints of the microdeletion in carrier foetuses and they identified a 128225-bp region deleted in the PEG3 domain, between base pairs 64320684 and 64448909 based on UMD 3.1 (Study II/Figure S1A).

The bovine genome assembly Btau4.2, together with a comparative map of corresponding regions in other species, indicates that the deletion eliminates the 3′ end of the non-coding *MIMT1* gene and part of the intergenic region between the *MIMT1* and *USP29* genes. The rest of the deleted BTA18 region contains several repetitive elements, such as retrotransposons, long terminal repeats and tandem repeats.

4.1.4 The *de novo* microdeletion is inherited by the progeny of affected bull (I and II)

The SNP genotyping data showed that the proband bull's sire was heterozygous in the deletion area, and thereby deemed to be free of the deletion. This result was expected since the proband bull had 33 half-brothers from the father side in AI, and none had any problems with abortions or stillbirths. Unfortunately, only 4–5 hair-straws with root bulbs were available from the mother dam of the proband bull. However, we succeeded in analyzing two SNPs in the deletion area and the dam tested heterozygous for one of the loci, and indicated that the dam was also free of the deletion.

The haplotype analyses indicated that some of the unaffected calves might carry the microdeletion. We used the sequence data to develop a simple, robust screen to identify affected animals by PCR amplification across deletion breakpoints (Study II/Figure S1B and C). We used the PCR screen to study 43 first-generation (F_1) and 25 second-generation (F_2) descendants of the proband bull. Four F_1 animals were found to carry the microdeletion, and all were female with normal reproductive capability. One young F_2 bull was also identified as a deletion carrier, showing that the deletion can be inherited by the next generations.

4.1.5 Microdeletion causes a loss of *MIMT1* expression (I)

Based on the localization of the deletion, it was expected that *MIMT1* expression would be disrupted in progeny that inherited the deletion chromosome from the proband bull. Furthermore, the deletion might affect the expression of the surrounding genes. Expression of *MIMT1*, *PEG3* and *USP29* genes was examined in foetal brain and cotyledon tissue (*pars fetalis*) from the eight foetuses (C1–C8). Two different-sized alternatively spliced *MIMT1* transcripts, 644 and 737 bp in length, were detected in control foetuses. While in foetuses that inherited the deletion from the proband bull the expression of *MIMT1* was absent indicating truncation of the gene (Figure 6A). We sequenced the corresponding RT-PCR products and found that the nucleotide sequence of exon 2 is alternatively spliced in MIMT1, confirming previously published findings (34). In case foetuses the expression of PEG3 was lower in brain and cotyledon tissue (Figure 6B), but expression of USP29 was lower only in brain tissue (Figure 6C). Both transcripts were detectable in both tissues of all foetuses.

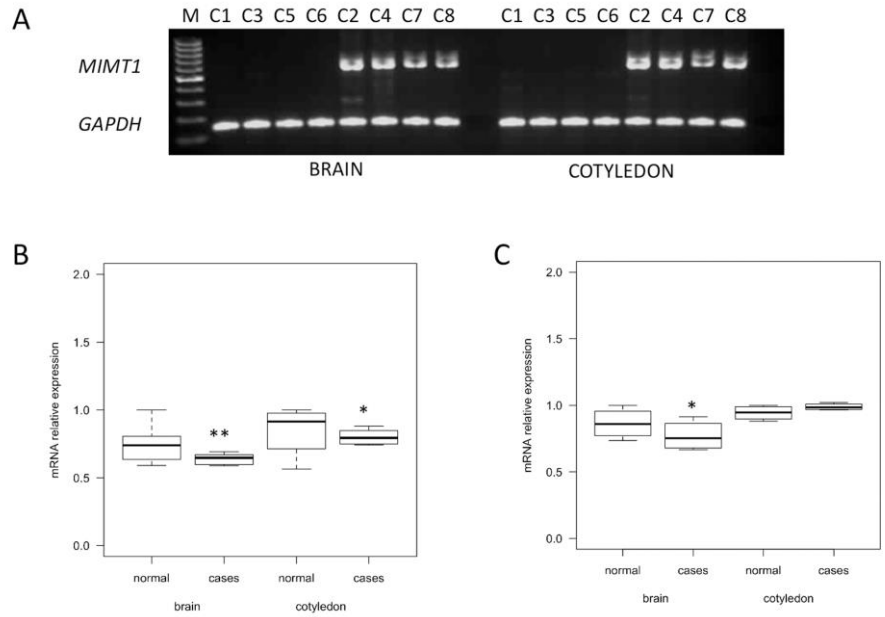

Figure 6. (A) The RT-PCR analysis of the *MIMT1* gene shows two different sized bands in foetuses without the microdeletion (C2, C4, C7, C8) due to the alternatively spliced exon 2. No product was detected in foetuses with the microdeletion (C1, C3, C5, C6). *GAPDH* was used for the confirmation of reverse transcription efficiency. (B and C) Box plots showing qPCR analysis of *PEG3* (B) and *USP29* (C) between carrier and non-carrier foetuses. qPCR expression is shown as a relative level of mRNA calculated with the $2^{(-\Delta\Delta ct)}$ formula and using *GAPDH* as an internal control gene. P-values: * <0.05, ** <0.01.

4.1.6 Microdeletion affects gene expression in foetal placenta (II)

All affected calves were undersized and because placental dysfunction is a primary cause of growth retardation the next goal was to study the gene expression levels in the placental tissues of foetuses. Samples of 14 new foetuses were collected and in total 22 foetal samples were tested for deletion and *MIMT1* expression as described earlier. Of these foetuses 11 carried the deletion and expressed truncated *MIMT1* (referred to as $MIMT1^{Del/WT}$) and 11 carried the wild type allele and expressed *MIMT1* ($MIMT1^{WT/WT}$). Samples were collected on 41–154 days of pregnancy and at this point there were no evident difference in foetal size or development.

To identify transcription differences between *MIMT1*$^{Del/WT}$ and *MIMT1*$^{WT/WT}$ samples of foetal cotyledon were analysed using custom gene expression arrays (SurePrint G3, Agilent). Six *MIMT1*$^{Del/WT}$ and four *MIMT1*$^{WT/WT}$ samples were analysed with microarrays. Datasets were processed, normalised and initially analysed with correlation heat maps to cluster the datasets of individual samples according to their pairwise correlations. Differentially expressed genes between *MIMT1*$^{Del/WT}$ and *MIMT1*$^{WT/WT}$ age-matched samples were identified with statistical analysis. Differential expressions with nominal P values of < 0.01 were found with over 430 probes but no significant probes were obtained after correction for multiple testing. Genes showing most marked differences in expression on the basis of fold differences and nominal P values were selected for further data analysis. A total of 140 genes showed increased and 108 genes showed reduced levels of expression in *MIMT1*$^{Del/WT}$ (Study II/Table S2). These genes were subjected to hierarchical cluster analysis and the heat map in Study II/Figure 2 reveals that *MIMT1*$^{WT/WT}$ samples formed a relatively homogeneous group with similar expression patterns. The *MIMT1*$^{Del/WT}$ samples were more heterogeneous, with sample C9 being the most similar to the control samples.

Nine genes were selected for qRT-PCR analysis on the basis of marked differential expression (*IL1RN*, *SMOC2*) or known importance in prenatal development (*IGF2*, *DLK1*, *MEST*, *AST1*, *PEG3*, *APEG3*, *H19*). Analysis was performed with 11 *MIMT1*$^{Del/WT}$ and 11 *MIMT1*$^{WT/WT}$ cotyledon samples, including those used for microarray study. All nine genes showed differences in expression very similar to those detected by microarray (Study II/Table 1).

Interestingly high expression of NPSR1 in four of the six MIMT1Del/WT cotyledon samples was found in microarray analysis. The finding was confirmed by the RT-PCR analysis of all 22 cotyledon samples. None of the MIMT1WT/WT foetuses expressed NPSR1 while seven of the MIMT1Del/WT foetuses did express the gene (Figure 7, Study II/Table 2).

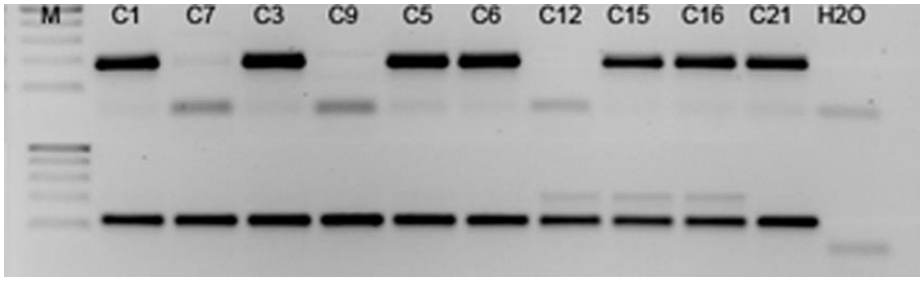

Figure 7. The RT-PCR analysis of *NPSR1* (upper bands) and *GAPDH* (lower bands) expression in *MIMT1*$^{Del/WT}$ cotyledons.

In the microarray analysis, the gene expression pattern of foetus C12 and, more strikingly, foetus C9 bore a closer resemblance to control samples than did the other *MIMT1*$^{Del/WT}$ foetuses. One clear difference between these and other *MIMT1*$^{Del/WT}$ samples is the lack of aberrant *NPSR1* expression. Excluding C9 and C12 from the microarray data comparison revealed more marked differences in several gene expressions (Study I/Table 1 and Table S2).

In humans, the *NPSR1* gene expresses two mRNA isoforms, A and B, which arise from alternative 3′ exon use (119). The *NPSR1A* expression was analyzed using RT-PCR with a specific primer pair (Study II/Figure S3) in cotyledon, brain, lung, kidney, heart, liver, and muscle samples of *MIMT1*$^{Del/WT}$ and *MIMT1*$^{WT/WT}$ foetuses and in the corresponding maternal caruncles for each foetus. The *NPSR1* mRNA expression was detected neither in foetal tissues, other than cotyledon, nor in the maternal caruncles.

The NPSR1A protein expression was studied in cotyledon samples of three *MIMT1*$^{Del/WT}$ foetuses and one *MIMT1*$^{WT/WT}$ foetus with Western blotting (Study II/Figure 4). Analysis revealed a ~50-kDa species, similar to that detected in human tissues (119), in *MIMT1*$^{Del/WT}$ foetuses, but not in the *MIMT1*$^{WT/WT}$ foetus.

4.1.7 Microdeletion results in altered global DNA methylation levels (II)

To establish if the deletion has an effect on methylation we analysed the global DNA methylation. The detected means in foetal cotyledon

and brain samples were 31.5 ± 6.82% and 71.2 ± 4.43 % respectively. Significantly greater global DNA methylation was found in *MIMT1*$^{Del/WT}$ cotyledons (35.08 ± 6.18 %) than in *MIMT1*$^{WT/WT}$ cotyledons (27.2 ± 4.91 %; P = 0.003) (Study II/Figure 1). Also an age-dependent but statistically non-significant decrease of DNA methylation in all samples was found. However, all *MIMT1*$^{Del/WT}$ cotyledons exhibited a higher level of DNA methylation than *MIMT1*$^{WT/WT}$ cotyledons of similar age. No differences in foetal brain DNA methylation were detected.

4.2 Inherited gonadal hypoplasia is associated with homozygosity of a chromosomal segment that has been duplicated and translocated from BTA6 to BTA29

4.2.1 Animals considered affected by gonadal hypoplasia

Unilateral gonadal hypoplasia was diagnosed in 11 cows and heifers with one extremely small ovary (Figure 8A). Usually the hypoplastic ovary was undetectable by palpation per rectum. Nine bulls and bull calves were considered to be affected by unilateral gonadal hypoplasia. Their testicles clearly differed in size, i.e. one testicle being more than twice as large as the other in young calves, and for bulls one testicle being more than three times larger than the other (Figure 8B). Bilateral gonadal hypoplasia was diagnosed only in one bull. Both of its testicles were very small and there were no sperm cells in the ejaculate. The histology showed only Sertoli cells and no spermatogonia in seminiferous tubules in the hypoplastic testicles.

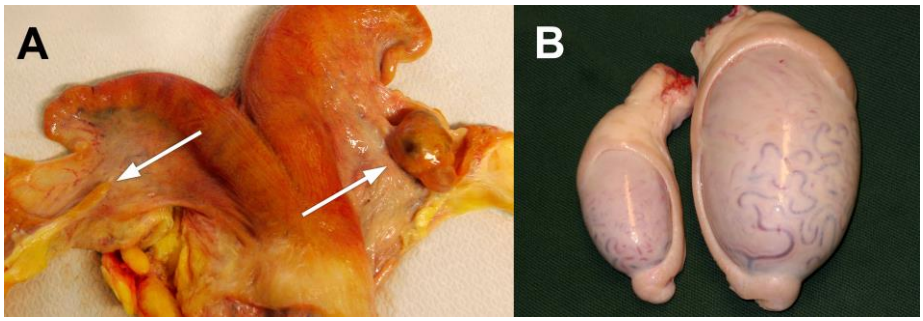

Figure 8. (A) Hypoplastic ovary (left arrow) and normal ovary (right arrow). (B) Hypoplastic testicle (left) and normal testicle (right).

4.2.2 The congenital gonadal hypoplasia maps to bovine chromosome 29

Gonadal hypoplasia and predominantly white coat colour in Northern Finncattle and Swedish Mountain cattle are probably associated, indicating pleiotropic effects of one gene (51, 120). Four different case-control designs were established (Study III/Table 1) to account for both the different coat colour patterns and the affection status of the animals. Allelic associations of these groups were performed and P-values below 7.71 x 10-8 (Bonferroni-corrected threshold for multiple testing) were considered to indicate significant association.

In the first GWAS (Study III/Table 1, Design A) 20 unaffected predominantly black and brown animals were compared with 21 affected predominantly white animals. The analysis revealed 18 significantly associated SNPs on BTA29 between 17.69 Mb and 20.50 Mb (Figure 9A, Study III/ Table S2). The SNP BovineHD2900005672 showed the most significant association (19,661,149 bp, P=2.19 x 10-13). Comparison of all unaffected animals (N=73) and affected animals (N=21) (Study III/Table 1, Design B) yielded five highly significantly associated SNPs on BTA29 and one highly significantly associated SNP on BTA23 at 52,435,290 bp, (P=9.57 x 10-9) (Figure 9B). There were no other significantly associated SNPs on BTA23, which indicated that this SNP was a false positive result. Moreover, technical problems with the SNP genotyping were assumed because re-genotyping the SNP on BTA23 using Sanger-sequencing revealed

discrepant genotypes compared with the genotypes obtained with the Illumina BovineHD Bead chip in some samples.

The third GWAS, comparing 40 unaffected and 21 affected animals all predominantly white (Study III/Table 1, Design C), revealed a suggestive association (Study III/Table S2). Similarly, GWAS in the fourth design (Study III/Table 1, Design D) with 40 unaffected predominantly white animals and 20 unaffected predominantly coloured animals revealed a suggested association (Study III/Table S2). However, after Bonferroni-correction the last two GWAS analyses indicated no significant association (Figure 9C and D).

Additionally, all four association analyses were repeated using a genotypic test and the results were in line with the allelic test. The genomic inflation factors for the groups ranged from 0.94 to 1.02 and indicated the absence of potentially spurious associations.

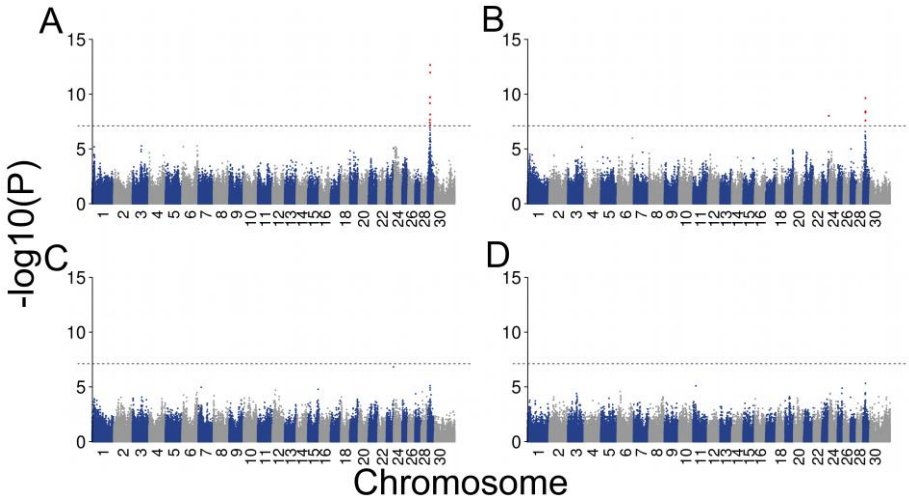

Figure 9. The Manhattan plots represent the $-\log_{10}(P)$ values of association for 647,971 SNPs in four different case-control designs (Study III/Table 1). The red dots represent significantly associated SNPs ($P < 7.71 \times 10^{-8}$).

4.2.3 Gonadal hypoplasia is associated with a chromosomal segment that has duplicated and translocated from BTA6 to BTA29

Predominantly white coat colour associated with gonadal hypoplasia in Northern Finncattle and Swedish Mountain cattle is likely due the colour-sided pattern that is determined by two translocated CNV segments in various cattle breeds (24). CNV analysis was carried out to investigate whether similar duplications are present in Northern Finncattle and the Swedish Mountain breed. A total of 2101 autosomal CNVs were identified in 94 animals.

The CNV segment of BTA6, extending from 71.60 Mb to 72.13 Mb and containing the *KIT* gene, was significantly associated with gonadal hypoplasia ($P=5.65 \times 10^{-6}$) (Study III/Figure 3, Figure 10A). The CNV was present in all affected animals and in 52 of 73 animals of the control group. To identify potential translocation loci, 38 animals with CN four and 21 animals with CN two were compared with GWAS. The CN status was determined based on the SNP array genotypes. We found a strong association on BTA29 (19,661,149 bp, $P = 1.66 \times 10^{-28}$), indicating a translocation of the BTA6 segment to BTA29 (Study III/Figure S3). The position agrees with the earlier results (24), thus suggesting that this translocation is Cs_{29}.

The CNV segment of BTA29, extending from 19.86 Mb to 20.32 Mb was identified in 14 affected and 44 unaffected animals (Figure 10B), but was not associated with gonadal hypoplasia (Study III/Figure S2, Tables S4 and S5) even thought the segment is in the immediate vicinity of the most significantly associated gonadal hypoplasia SNPs. Strong association on BTA6 (72,198,048 bp, $P=1.69 \times 10^{-14}$, Study III/Figure S4) was found after the GWAS comparison of 58 animals with CNV of BTA29 and 36 animals without CNV of BTA29. This indicates the translocation of CNV segment on BTA29 to BTA6 and denotes translocation as Cs_6 because the position is the same as in the earlier study (24).

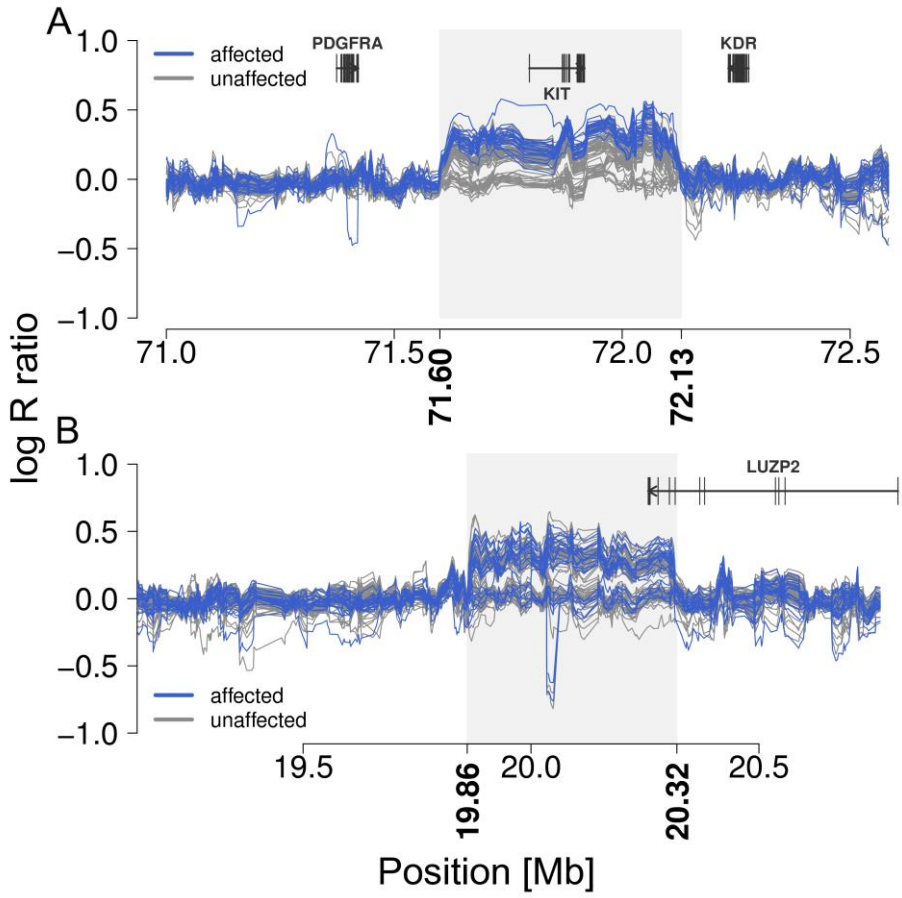

Figure 10. The figures display 5-SNP-sliding-window log R ratios in 75 unaffected (grey) and 21 affected (blue) animals on BTA6 (A) and BTA29 (B). The grey shaded box represents the extent of two CNV segments.

4.2.4 Validation of the translocated copies reveals that the underlying mechanism of gonadal hypoplasia is the homozygosity of Cs_{29}

The discovered translocations were confirmed with cytogenetic studies and qPCR. Three Northern Finncattle animals with different Cs_{29} status were selected for the cytogenetic studies. The animal with two Cs_{29} alleles was affected by gonadal hypoplasia. A Western Finncattle and an Eastern Finncattle animal were also included in the cytogenetic

analysis. Traditionally, Western Finncattle is a solid brown breed and Eastern Finncattle is a brown breed with the colour-sided phenotype and these breeds are not affected by gonadal hypoplasia. Results are shown in Study II/Figure S5. The Cs_{29} alleles of nine affected and 21 unaffected animals were analysed using qPCR. The results obtained with cytogenetic studies and qPCR agree with the CNV results based on the SNP array genotypes. However, qPCR could not distinguish between heterozygous and homozygous animals.

Conventional PCR with most primers, according to the earlier study (24), was also used for the validation of the translocations. Analysis included all SNP genotyped animals and also two unaffected animals that were omitted from the GWAS analyses because of the failure of genotyping. The colour of these two animals was predominantly brown and white. All 21 affected animals were homozygous for Cs_{29} but only 15 of them had the Cs_6 allele. Of the 75 unaffected animals, 18 were homozygous and 37 were heterozygous for the Cs_{29} allele and 20 had wild type alleles on BTA29. The Cs_6 allele was detected from 44 control animals (Table 2).

Table 2. The PCR analysis of Cs alleles, including 96 animals of Northern Finncattle and Swedish Mountain cattle breeds. The animals are divided according to Cs_{29} and Cs_6 alleles. Heterozygous and homozygous carriers of the Cs_6 allele could not be distinguished.

	Affected			**Unaffected**		
	+/+	+/Cs_{29}	Cs_{29}/Cs_{29}	+/+	+/Cs_{29}	Cs_{29}/Cs_{29}
+/+	0	0	6	10	17	4
Cs_6/-	0	0	15	10	20	14
total	0	0	21	20	37	18

The ratio of the Cs_6 allele is slightly higher in affected (71%) than in unaffected animals (59%) but more importantly the occurrence of Cs_6 (78%) in control animals that were homozygous for the Cs_{29} allele is almost the same as in affected animals (Table 2). As with the case animals, all homozygous control animals were predominantly white. These results suggest that the homozygous Cs_{29} allele has pleiotropic effects on both coat colour and gonadal hypoplasia, but the heterozygous Cs_{29} allele and the Cs_6 allele affect only the coat colour.

All conventional PCR results are in line with the CNV results based on array genotypes, cytogenetic analyses and qPCR. Additionally, we determined the DNA sequence of all PCR products amplified from the selected 12 animals (results not shown). The sequences were comparable with those from earlier research (24).

4.2.5 Comparison of the case and control animals homozygous for the Cs_{29} allele revealed no association

Given that all affected animals were homozygous for the Cs_{29} allele, we performed an additional conditional GWAS where the affected animals were compared with 18 control animals that were homozygous for the Cs_{29} allele. The only suggestion of an associated allele was the SNP on BTA23 at 52,435,290 bp, which was shown to have discrepant genotypes (Study III/Figure S6).

4.3 In frame exon skipping in *UBE3B* is associated with PIRM syndrome in Ayrshire cattle (IV)

4.3.1 PIRM syndrome in the Ayrshire population

Farmers and breeding advisers reported an increasing number of calves with various developmental defects in the Finnish Ayrshire population between 2011 and 2014. The most obvious and prevalent symptom was ptosis, but some affected calves also suffered from feeding problems, minor structural changes of the head and muscular hypotonia. Many of these affected calves failed to thrive and died at a very young age if not euthanized before. Surviving calves required special care during the neonatal period and later showed growth retardation. Usually farmers culled affected animals before breeding. Breeders also reported learning difficulties, like problems learning how to use feeding buckets, indicating intellectual disability. The phenotype has been defined as PIRM syndrome according to its typical features: **p**tosis, **i**ntellectual disability, **r**etarded growth and mortality (Figure 11, Study IV/Table S1). Both sexes were equally affected and the analysis

of pedigree records of affected animals and their close relatives indicated an autosomal recessive mode of inheritance.

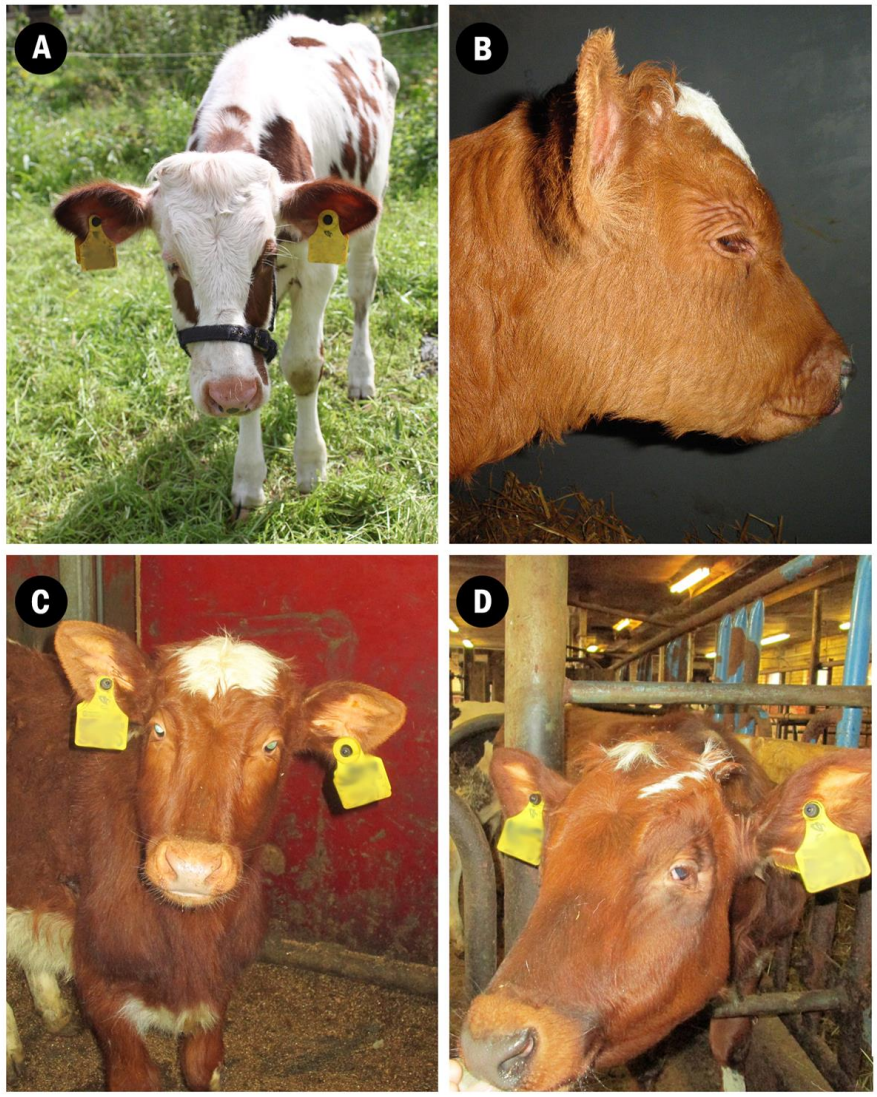

Figure 11. The most obvious and prevalent feature in PIRM syndrome is ptosis. Abnormally large upper eyelids characteristically make affected animals look as if they are constantly sleepy (A-D). Affected animals suffered also from hypotonia (A and B) or were smaller than other age-matched calves (C and D). In addition, owners reported problems with learning and dependency on special care. All animals were euthanized soon after the pictures were taken.

4.3.2 PIRM syndrome maps to a 713 kb segment on bovine chromosome 17

To identify the genomic region associated with PIRM syndrome, nine affected, 37 unaffected half-siblings and their AI sire were genotyped with a bovine high-density genotyping array. Genotypes for 623,881 SNPs were phased using the Beagle's hidden Markov model based algorithm and obtained haplotypes were then used in a genome-wide association study. Strong association on bovine chromosome 17 (Figure 12A) was found with a sliding window-based approach used to compare haplotype frequency in cases and controls. The most significant association ($P = 1.55 \times 10^{-9}$) resulted from four adjacent haplotype windows located between 65,659,074 bp and 65,981,740 bp.

To narrow down the associated region, the genotypes of affected animals were screened for segments of homozygosity. Extended homozygosity of 713 kb (65,645,831 bp - 66,358,629 bp) was present in all affected animals while none of the unaffected animals showed homozygosity, supporting the supposed recessive pattern of inheritance. The risk haplotype encompasses 14 genes (Figure 12B-C, Study IV/Table S2).

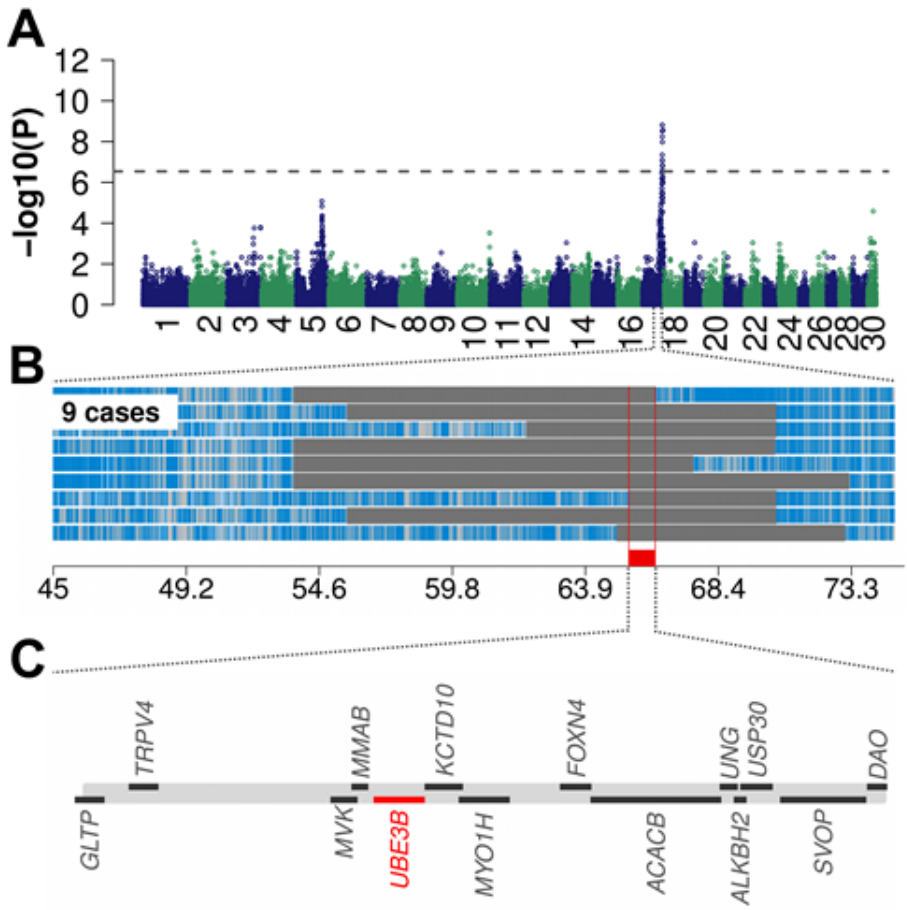

Figure 12. (A) Association of the affection status in nine affected and 37 unaffected Ayrshire half-siblings. P-values were obtained by calculating Fisher exact tests of allelic association. (B) Autozygosity mapping in nine affected animals. Blue and pale blue represent homozygous genotypes and heterozygous genotypes are displayed in light grey. The solid grey bars represent the segments of extended homozygosity in nine affected animals. The red bar indicates the common segment of homozygosity. (C) The shared segment of homozygosity encompasses 14 transcripts, among them *UBE3B*.

4.3.3 A synonymous substitution in *UBE3B* perfectly coincides with PIRM syndrome

For locating the underlying mutation responsible for PIRM, the whole genomes of an obligate carrier and one of its affected progeny were sequenced to average read depths of 10.4 and 10.1. No large structural elements in the PIRM-associated region were found with the sequence depth analysis (Study IV/Figure S1). Sequences of affected and carrier animals were compared with data from 43 previously sequenced Fleckvieh animals (121). The mutation causing the PIRM syndrome should not segregate among the sequenced Fleckvieh animals because there is no evidence that PIRM syndrome occurs in breeds other than Ayrshire. Multi-sample variant calling in 713 kb extended homozygosity segment yielded genotypes for 1684 SNPs and 141 InDels (Study IV/File S1). Among these, only four SNPs (Study IV/Table 1) were compatible with recessive inheritance that is homozygous for the reference allele in 43 Fleckvieh animals, heterozygous in the carrier bull and homozygous for the non-reference allele in the affected animal.

Based on annotation of the UMD3.1 bovine genome assembly (101) one of the four compatible variants was intergenic, two were located in intronic regions of *TRPV4* and *UBE3B*, and one variant was located in the coding region of *UBE3B* (Study IV/Table 1). However, two of the four variants (rs440561578 and rs467377722) segregated also among 191 non-Fleckvieh animals that were sequenced in the context of the 1000 bull genomes project (22), indicating that only the coding and an intronic variant of the *UBE3B* gene segregates with the PIRM syndrome.

Re-annotation of the bovine *UBE3B* gene showed that it consists of 32 exons, of which exons 1 to 6 are non-coding (Figure 13A, Study IV/Table S3). The variant in coding region of *UBE3B* (rs475678587) is a G > A polymorphism in the third base of codon 692 that causes a synonymous substitution, p.E692E in exon 23 of *UBE3B*.

The rs475678587 polymorphism was most likely the causative mutation for PIRM and therefore validated by Sanger sequencing in the carrier bull, its nine affected and 37 unaffected descendants. In addition, nine new cases and two control animals were also genotyped with a KASP genotyping assay. All 18 affected animals were

homozygous for the rs475678587 A variant, whereas the unaffected animals were either heterozygous or wild type. The suspected carrier bull and 23 of its descendants were heterozygous.

4.3.4 Mutation screening in Ayrshire AI bulls indicates a high carrier frequency

The frequency of the rs475678587 A variant was analysed with KASP reagents in 129 Ayrshire AI bulls that are in use or have recently been used in Finland. A carrier frequency of 17.1% was established, generating one affected calf per 138 offspring if random mating is assumed. From tested bulls, 29 had a known haplotype status for AH1, which was perfectly associated with the rs475678587 A mutation in this cohort. All 11 bulls that were carriers of the rs475678587 A mutation carry AH1 and 18 wild type bulls did not have AH1.

4.3.5 The rs475678587 A substitution affects splicing of *UBE3B*

The rs475678587 G > A polymorphism is located at the very last nucleotide in the junction between exon 23 and intron 23 (Figure 13B, Study IV/Table S3) and could therefore affect RNA splicing. The effect of the mutation was investigated using reverse transcription PCR and two primer pairs in samples of cerebral cortex, tectum, hippocampus, cerebellum, lung, liver, heart, kidney, spleen, and ovary from three affected and two unaffected animals. Primer pair, which flanked the exon 23 (Study IV/Table S4), resulted in the amplification of two fragments from the affected animals and only one fragment from the unaffected animals (Figure 14A). Sequencing revealed in-frame exon 23 skipping in the smaller PCR product and the other fragment of expected size corresponded to the reference sequence (UMD3.1, (92)) with the exception of the rs475678587 A variant in affected animals.

The *UBE3B* gene was expressed in all examined tissues and the altered splicing pattern was detected in all tissues of affected animals, excluding tissue specificity. The intensity of the amplicons obtained with primer pair flanking the exon 23 suggested lower relative

expression levels of *UBE3B* in affected animals, but results obtained with primer pair flanking exons 27–29 (Study IVI/Table S4) showed no difference between unaffected and affected animals (Figure 14B).

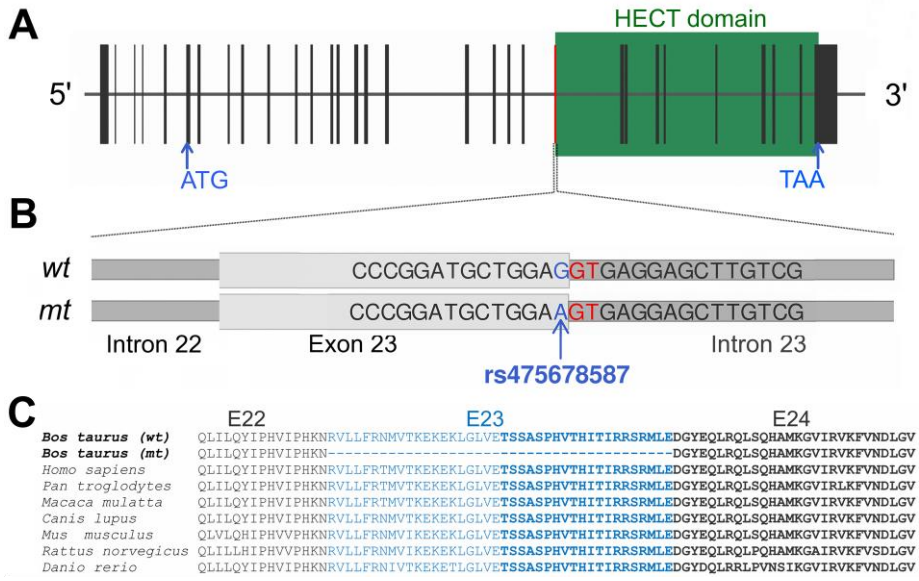

Figure 13. (A) The genomic structure of bovine *UBE3B*, which consist of 32 exons (vertical bars) and its translation starts in exon 7. The position of the HECT-domain was determined according to (86). (B) Wildtype (wt) and mutant (mt) sequence of exon 23 (to improve readability, only part of the genomic sequence is shown). The rs475678587 A mutation affects the very last nucleotide of exon 23 within the highly conserved HECT-domain. (C) The multi-species alignment of the UBE3B protein sequence. Alternating colour indicates different exons (22–24) and bold type indicates the initiation of the HECT domain.

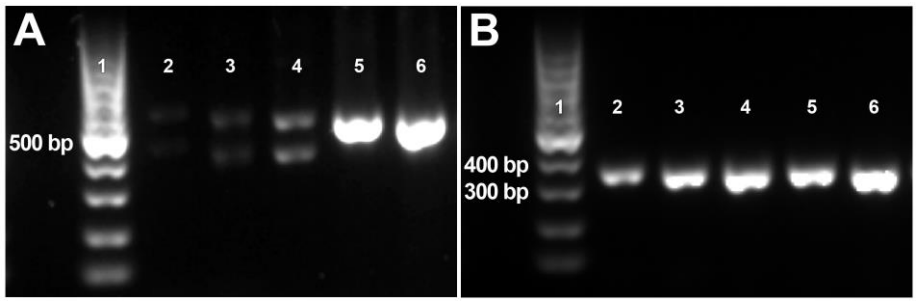

Figure 14. (A) The agarose gel electrophoresis of the RT-PCR products reveals that in addition to the normal band, affected animals also express a smaller fragment with primers that flanked the exon 23. (B) Amplicons obtained with primers that flanked exon 27–29 show no difference between affected and unaffected animals. Lane 1 is a molecular weight marker. Lanes 2–4 are cerebellum samples of PIRM affected animals and lanes 5 and 6 are unaffected animal samples.

4.3.6 *In silico* analysis predicts in-frame deletion of 40 amino acids

The amino acid alignment of the normal and mutated proteins showed that loss of exon 23 results in an in-frame deletion of 40 amino acids among residues 652–692. Twenty amino acids are deleted from the E2 binding subdomain of conserved HECT in UBE3B (Figure 3). The three-dimensional modelling of the mutated HECT domain revealed lack of one α-helix structure and considerable structural differences compared with the normal HECT domain (Study IV/Figure S2).

5 DISCUSSION

5.1 IUGR, late-term abortions and stillbirths in cattle are caused by a microdeletion in the imprinted PEG3 domain

The high mortality (almost 50%) and very small size of affected calves sired by one proband bull are caused by a microdeletion in the imprinted PEG3 domain. The deletion truncates the maternally imprinted and paternally expressed *MIMT1* gene. The small size of affected calves and truncation of a paternally imprinted gene is in line with conflict theory and observations that highlight the importance of paternally expressed genes for growth (27, 29).

The crossbreeding of the proband bull with Holstein females did not have an impact on offspring survival, which is consistent with the imprinted model of inheritance. The observed mortality rate of the proband's offspring was 42.6%, which suggests that about 85% of the $MIMT1^{Del/WT}$ calves die. The semi-lethality of the microdeletion was confirmed with findings that the tested progeny of the proband bull's four F1 generation and one F2 generation descendants were mutation carriers. The semi-lethality might be caused by an incomplete silencing of maternally imprinted alleles or by the reconstructive effects of other genes.

The imprinted model of inheritance also indicates that the microdeletion should be lethal only when inherited from the sire. This was true in tested F1 carrier animals, which were all females and had no reported problems with breeding. Furthermore, the breeding value estimate of the proband bull's stillbirth as a grandsire was 125 in (the average of the breed is 100 and standard deviation is 10). The estimate would not be so good if 15% of the daughters had problems with stillbirth. The F2 descendant was a young bull that was not used for breeding. According to the imprinted inheritance model, about half of pregnancies sired by carrier bulls should end in stillbirth as with the proband bull. The results indicate that both parents of the proband bull were free from the mutation, suggesting that the mutation could be *de*

novo in the proband bull. However, germ-line mosaicism in one of the parents could not be ruled out.

The *PEG3* knockout mouse models have shown the importance of *PEG3* gene function for normal reproduction, maternal care, growth and perinatal survival (122-125). Interestingly in mouse, the deletion of YY1 binding sites (part of the PEG3 DMR) caused imprinting and expression changes when the knockout allele was inherited from the father. Also these offspring showed increased embryonal lethality and reduced weight. No homozygote pups were born, indicating lethality during the gestation period and suggesting that at least one wild type allele is needed for the viable mouse pup (35). These earlier results, together with our results, underscore the importance of the PEG3 domain to growth and survival in mammals.

The *MIMT1* expresses lncRNA but its function in biochemical pathways remains unknown. The lncRNA transcripts regulate genome function and they can utilize both *cis* and *trans* acting mechanisms. In imprinted domains lncRNA genes have been shown to control the expression of the other genes of the domain (30, 31). Furthermore, the lncRNAs, with bidirectional promoters, might affect the function of the promoter and on that way change the expression of the gene transcribed in the opposite direction (126). Besides the truncation of *MIMT1,* the microdeletion removes part of the well-conserved middle region of the PEG3 domain that contains several ECRs. These ECRs are most likely putative enhancers and may also play a major role in gene expression of the domain (39, 41). Thus the microdeletion causes severe consequences that can be induced by the truncation of *MIMT1* or removal of ECRs or their joint effect.

The effects of the microdeletion on the neighbouring genes *PEG3* and *USP29* were analysed in brain and cotyledon samples. The expression levels appeared to be lower in four $MIMT1^{Del/WT}$ foetuses than in four $MIMT1^{WT/WT}$ foetuses. Furthermore, the microarray and qPCR analyses in a larger cohort of cotyledon samples showed the lowered expression of several genes of the PEG3 domain region (*AST1, PEG3, APEG3, LOC508098* and *MGC157368*) in $MIMT1^{Del/WT}$ foetuses. These results indicate that the microdeletion may have changed the regulatory mechanisms of these genes, but changes in imprinted expression were not studied.

The small size of affected calves indicates an intrauterine growth restriction (IUGR), which is often associated with the abnormal function of the placenta. Furthermore, several imprinted genes, including genes in the PEG3 domain, are expressed in the placenta, which is a key regulatory site for conflicting parental interests (127-129). The microarray expression analysis of cotyledon samples revealed differences in hundreds of genes involved in several physiological and biochemical functions. The genes with differential expression, and whose known functions were likely to affect prenatal development, can be grouped into two main categories: 1) imprinted genes (for example *IGF2*, *DLK1*, *MEST*, *PEG3*, and *H19)* (123, 125, 127, 130, 131) and 2) genes associated with blood circulation and inflammation (for example *NPSR1*, *IL1RN*, *NOS3* and *IL4R*) (132-136). Moreover, several genes were found with an unknown role in the placenta, but which showed strongly increased or decreased expression (for example *SMOC2*, *ZDHHC22*, *PRSS2* and *TKDP3*).

Possibly the most interesting discovery was the expression of *NPSR1* (Neuropeptide S receptor 1) gene in seven out of eleven $MIMT1^{Del/WT}$ cotyledons, while none of the $MIMT1^{WT/WT}$ cotyledons expressed the *NPSR1*. Expression of *NPSR1* was not found earlier in the placenta even though it is expressed in several cell types (137). *NPSR*1 has been associated with the immune responses, especially in asthma and allergy, neuroendocrine stress response and cell proliferation (138-140), but the function of *NPSR1* in the placenta remains to be discovered. Interestingly, in the microarray analysis *NPSR1* was not expressed in foetuses C9 and C12, which showed closer resemblance to control samples than to the other $MIMT1^{Del/WT}$ foetuses. As mentioned earlier, the discovered microdeletion is semi-lethal and it is tempting to hypothesize that expression of *NPSR1* is associated with the death or survival of mutation carrier foetuses as a cause or a consequence.

The aberrant DNA methylation in the placenta is also associated with IUGR (141, 142). The global methylation analyses showed a 7% increase in $MIMT1^{Del/WT}$ placenta samples when comparing with $MIMT1^{WT/WT}$. These methylation changes might be associated with the revealed gene expression changes. Particularly the methylation changes of imprinted genes in the placenta are linked to foetal growth

and development (142). The global methylation differences were found before any obvious changes in growth or pathology were apparent in *MIMT1*$^{Del/WT}$ foetuses, indicating a causative role of the methylation for intrauterine growth restriction and lethality.

In conclusion, we have shown that the microdeletion in the imprinted PEG3 domain causes IUGR that leads to late abortions and stillbirth. The deletion truncates the paternally expressed *MIMT1* and probably removes ECRs that are suggested to be enhancers of surrounding genes. The gene expression and methylation analyses in the placenta indicate that the deleted region is associated with the complex network of regulatory elements. To our knowledge, this was the first time that mutation in the imprinted domain has been associated with an inherited bovine developmental disorder.

5.2 Inherited gonadal hypoplasia in predominantly white animals of Northern Finncattle and Swedish Mountain cattle is associated with the homozygosity of the *Cs*$_{29}$ allele

The gonadal hypoplasia of Northern Finncattle and Swedish Mountain cattle is associated with homozygosity of the *Cs*$_{29}$ allele, which is a result of a complex chromosomal rearrangement, including a ~500kb duplicated segment from BTA6 which is translocated to BTA29. The allele includes the entire coding sequence of the *KIT* gene, indicating that affected animals have four functional *KIT* genes (24).

The gonadal hypoplasia in Northern Finncattle and Swedish Mountain cattle is strongly associated with the predominantly white coat colour, which is linked to colour sidedness (51, 56, 57). This coat colour pattern is partly conferred by the *Cs*$_{29}$ allele (24), indicating that the gonadal hypoplasia and white coat colour phenotypes are both associated with the *Cs*$_{29}$ allele. This is supported by the GWAS results that showed no association between predominantly white affected and unaffected animals. Moreover, the association signal was stronger when affected animals were compared with unaffected coloured animals than when affected animals were compared with all unaffected animals. The latter group included several predominantly white animals that carried the *Cs*$_{29}$ allele, explaining the weaker signal and

underscoring the pleiotropic effect of the CS_{29} allele. In the light of this, the result that the unaffected predominantly white animals and unaffected predominantly coloured animals showed no association was surprising. Most probably the explanation for this result is that some of the coloured animals were colour-sided, harbouring the Cs_{29} allele, and some of the white animals did not carry the Cs_{29} allele. Thus, the GWAS results indicate stronger association between the Cs_{29} allele and gonadal hypoplasia than between the Cs_{29} allele and predominantly white coat colour. Furthermore, this strongly suggests that the Cs_{29} allele is associated with gonadal hypoplasia and not only reflects different coat colour patterns.

The colour sidedness is also associated with the Cs_6 allele (23, 24), but our GWAS and CNV results indicate that it is not associated with gonadal hypoplasia. Furthermore, PCR results showed that the Cs_6 allele is only slightly enriched in affected animals when compared with all unaffected animals. More importantly, the frequency of the Cs_6 allele was almost the same in case animals as in control animals that were homozygous for the Cs_{29} allele.

The effect of the *Cs* alleles on coat colour is presumably caused by the misregulation of the *KIT* gene that leads to impaired migration, proliferation and survival of melanocyte precursors (24). Mutations in the *KIT* gene cause white colour also in other mammals like pig (OMIA 000209-9825, 001743-9825 and 001216-9825), cat (OMIA 000214-9685, 000209-9685) and horse (OMIA 000209-9796) (6), but there are no reports of associations with gonadal hypoplasia. On the other hand, several mouse models harbouring *KIT* mutations exhibit pleiotropic affects, causing reduced fertility and white colouring. Many of these mice mutants suffer also from anaemia, tumour formation and deafness (MGI:96677) (66). Yet the gonadal hypoplasia types in mouse studies do not resemble the gonadal hypoplasia of Northern Finncattle and Swedish Mountain cattle. Moreover, we are not aware that the affected animals in our study had any other symptoms besides gonadal hypoplasia.

A similar type of gonadal hypoplasia as in the Northern Finncattle and Swedish Mountain cattle was reported in the Nguni breed (*Bos indicus* x *Bos taurus* cross) (143-145). Interestingly, this breed also presents the colour-sided phenotype. The colour sidedness is found in

several breeds and *Cs* alleles are suggested to account for most colour sidedness in cattle (24). Moreover, in White Galloway cattle and White Park cattle the predominantly white coat colour was attributed to the Cs_{29} allele (58). The same link was shown in our study where only five animals out of 62 predominantly white animals did not have the Cs_{29} allele. On the other hand, in yaks Cs_{6} was more frequent in the white animals than the Cs_{29} allele, which could be due to the small cohort or breed differences (23).

The Cs_{29} allele has been found in several breeds (23, 24, 58). The generalization of the *Cs* alleles may be due to the selection for coat colour patterns because they may constitute one of the most important phenotypic traits for breed identity. However, hereditary gonadal hypoplasia has been reported only in Northern Finncattle, Swedish Mountain cattle and Nguni cattle. This might be because the proportion of animals homozygous for the Cs_{29} allele is too small in several colour-sided breeds for this type of recessive gonadal hypoplasia with incomplete penetrance to be diagnosed. Another hypothesis is that Swedish Mountain cattle and Northern Finncattle have other breed-specific genomic changes in addition to the Cs_{29} allele. This conclusion implies that Nguni might have different mutations that cause the gonadal hypoplasia.

Importantly, 18 unaffected animals in our study were homozygous for the Cs_{29} allele. This most probably is because of incomplete penetrance, which was suggested to be a partly an inherited feature. The inherited penetrance could imply additive mutation or mutations, but the GWAS and CNV studies showed no other reliable association between the cases and controls than the association for the Cs_{29} allele. However, an additional mutation might remain undetected because of the small sample size in the present study or because of the location of the mutation. If the mutation is within the translocated area or in a corresponding area in the original chromosome, the mapping with GWAS might not work.

In conclusion, the hereditary form of gonadal hypoplasia in Northern Finncattle and Swedish Mountain cattle is associated with homozygosity for the Cs_{29} allele. This is in line with the postulated mode of recessive inheritance (52) and highlights gene dosage as a significant risk factor for the defect. However, there is a need for

further investigations to understand all the mechanisms causing the disorder.

5.3 Association of PIRM and AH1 haplotype with the exon skipping mutation in *UBE3B*

The newly discovered developmental syndrome PIRM in Ayrshire cattle is associated with an exon skipping mutation in the *UBE3B* gene. This defect is congenital and manifested by a complex disease phenotype. Exon skipping results in a partial truncation of the UBE3B protein, likely compromising its function.

The *UBE3B* impairments in human are associated with KOS (75-78) and with autism in one family (146). The KOS patients show a variety of severe symptoms, including developmental delay with intellectual disability, breathing and feeding difficulties, hearing impairment, hypotonia, retarded growth and hypocholesterolemia. The KOS patients also have characteristic craniofacial dysmorphisms like ptosis, blepharophimosis, microcephaly and telecanthus that may result from maldevelopment of first and second branchial arches (75-78). Mice engineered to lack *UBE3B* expression ($UBE3B^{-/-}$) displayed mild hearing impairment, diminished body weight and size, reduced grip strength, which may reflect central or muscular hypotonia, and significant reduction of total cholesterol and lathosterol. Abnormalities in brain and eye structures were also found. Additionally, increased embryonal and perinatal lethality was reported for $UBE3B^{-/-}$ mice (76). Symptoms caused by *UBE3B* loss-of-function in human and mice are strikingly similar to pathological signs observed in PIRM Ayrshire cattle.

As with most KOS patients the *UBE3B* bovine mutation most likely causes malfunction of the HECT domain (75-78). The bioinformatic analysis predicted that the mutated UBE3B protein lacks 40 amino acids, of which 20 are included in the highly conserved and catalytic HECT domain. Interestingly both wild type and mutated transcripts are expressed in PIRM-affected animals, indicating that also normal protein is produced. Expression of two transcripts is most likely due to ineffective splicing. Other developmental disorders with similar in-frame exon skipping mutations have been found, for

example, anophthalmia/microphthalmia (147), cholesteryl ester storage disease (148) and dystrophic epidermolysis bullosa (149). *UBE3B* mutations in heterozygous form do not cause any clinically detectable symptoms in cattle, which accords with studies on human and mouse (75-78, 146).

In cows, *UBE3B* is expressed in several tissues, which is in line with the human (86, 150, 151) and mouse (76, 152) studies. However, the functional role of UBE3B in different biochemical pathways is still unknown. In the cell, the UBE3B protein is mainly located in the nucleus and additionally in the mitochondria (151) where it may have a role in mitochondrial protein turnover (153). In chicken, noise trauma was found to cause up-regulation of *UBE3B* mRNA expression in damaged areas of the inner ear (150). Furthermore, the *UBE3B* orthologue OXI-1 in *C. elegans* is associated with oxidative stress-response (76, 154) indicating association of *UBE3B* and protein degradation under stress conditions. Increased sensitivity toward oxidative stress might be associated with neuronal dysfunction (76). The *UBE3B* gene has also been associated with the depolarization-regulated gene transcription of neuron cultures, which is related to synaptic plasticity (146). The observed hypocholesterolemia in KOS patients and UBE3B$^{-/-}$ mice, together with the association of UBE3B and total cholesterol in plasma, indicates that UBE3B might have a role in cholesterol synthesis (76, 155). The blood-brain barrier prevents the brain from utilizing the circulating cholesterol, making brains particularly vulnerable to the deficiency of cholesterol synthesis (76, 156).

The Ayrshire haplotype 1 (AH1) is located in BTA17 (14) and encompasses the rs475678587 polymorphism. Moreover, in our cohort AH1 was perfectly associated with the nucleotide substitution A, which was also associated with the PIRM syndrome. This might indicate that the observed lack of AH1 homozygous animals in the earlier study might be linked to the severe symptoms of PIRM. The AH1 was related to reduced fertility that could imply embryonic losses, which agrees with the elevated embryonic lethality in the UBE3B$^{-/-}$ mouse model (76). Thus this could imply that reduced fertility is also one of symptoms caused by the *UBE3B* mutation in cattle. The effect of AH1 on the stillbirth rate could not be determined.

The estimated frequency of the AH1 haplotype was 26.1% in the North America Ayrshire herd (14), even higher than the 17.1% frequency in our AI bull cohort. The high and steady prevalence of the AH1 in the North American population may be the cause of genetic drift resulting from heavy use of the carrier ancestor Selwood Betty's Commander (14). However, the AH1 or mutated *UBE3B* may also be linked with desired traits that have caused the selection benefit for carrier animals.

The similar location to the position of AH1 was studied in the Nordic Red cattle population, which includes Danish Red, Swedish Red and Finnish Ayrshire breeds. The study included about 38,000 animals and one haplotype was found that was present in homozygous form only in five animals (no phenotype information about these animals was given). The frequency of heterozygote carriers was 6.8%, indicating one affected calf per 865 offspring if random mating is assumed. The haplotype found reduced the breeding value estimate for calving by 3.8 points (with the standard deviation of breeding value estimates being 10), but the effect on the fertility (non-return rate) was not studied. The association of the *UBE3B* mutation with the haplotype in Nordic Red cattle has not been studied (157).

In conclusion, our study shows that mutation in *UBE3B* is associated with PIRM syndrome in cattle. PIRM resembles human KOS and shares features with the *UBE3B* knockout mouse models. Our results support the importance of UBE3B protein for normal development in mammals. The high prevalence of the *UBE3B* mutation in Ayrshire AI bulls and the possible link between PIRM and AH1 is a matter of considerable concern for cattle breeders and ranchers in several countries. Our study does however provide an avenue for further investigation and control of PIRM.

5.4 Practical and theoretical implications

The results from this study have practical and theoretical implications in veterinary medicine and beyond. The causative mutations of the described inherited defects are now relatively easy to test for. The results can be utilized to avoid risky matings and culling carriers. In addition, DNA tests can also be used for veterinary diagnostics.

Particularly with PIRM syndrome the high carrier frequency among the tested AI bulls and severe symptoms of the disease highlight the importance of controlling the defect. Furthermore, the possible association with subfertility, which is an endemic adverse trait in modern cattle breeds (158, 159), also underlines the importance of reducing the mutation frequency. However, the possible connection of AH1 and the *UBE3B* mutation to desired traits should be studied and results applied to breeding plans. In the report of the United States Department of Agriculture, the *UBE3B* mutation is listed as causal for the infertility effect of haplotype AH1 (160). In light of this, it is possible to use the AH1 haplotype as a genotype marker for the carrier analysis of the *UBE3B* mutation.

We also showed that the mutation that causes stillbirth in Ayrshire cattle can be inherited and most likely the deletion causes no symptoms when inherited from the mother. This indicates that the mutation in the PEG3 domain can spread unnoticed without any adverse effect as long as carrier bulls are not used for breeding. Hence if male descendants of the proband bull are to be used for breeding they should be first DNA tested. Also the possibility of mosaicism in the parents of the proband bull should be taken into account when their male descendants are chosen for breeding purposes. The substantial level of stillbirths in the first generation of the proband bull, similarly to defects caused by the dominant mutation, emphasizes the problematic nature of excessive use of young AI bulls without progeny testing.

The prevalence of the Cs_{29} allele in our study cohort, and that the allele is not the only causative factor of gonadal hypoplasia, does not support elimination of carriers from this endangered breed. Furthermore, the cardinal left-sided manifestation of the defect may not compromise the fertility of the affected animals. Possible effects on other traits have not been studied. The predominantly white coat colour is still favoured in breeding Northern Finncattle. However, the association of white coat colour and gonadal hypoplasia raises a question concerning the importance of the exterior traits of the breed.

Genes with documented biological function related to valuable traits can also serve as the basis of candidate gene approaches for quantitative trait loci (QTL) studies, which may lead to the discovery

of new markers for breeding. For example, the normal function of *UBE3B* is needed for postnatal growth, making it a possible target gene for growth trait studies. Also the *KIT* gene is strongly correlated with fertility and coat colour traits. Furthermore, the association of imprinted genes with development, growth, metabolism and maternal care makes them an interesting target for QTL studies in cattle. Thus several associations between imprinted genes and QTs have been found (161) including the relation of the PEG3 domain genes *ZIM2* and *PEG3* with traits directing neonatal development and survival (162). Our study clearly links the mid-PEG3 domain and *MIMT1* with the growth and survival of the foetus, making it an attractive genomic area for QTL studies. It is likely that the imprinted genes represent an important reservoir for future genetic improvement of cattle.

In addition to breeding, mapping the causative mutations of inherited defects in cattle provides new insights on genotype-phenotype associations. The comparative studies across species are relevant for biomedical research and may lead to discoveries that could have an impact on human and veterinary medicine. The *KIT* and *UBE3B* genes and the PEG3 domain are essential for normal organogenesis and growth during the foetal period, indicating their relevance to developmental biology research. The studies related to the microdeletion causing IUGR and stillbirths in Ayrshires are excellent examples of how naturally occurring mutations in livestock can be used as models for basic molecular studies. To date three publications reporting expression and methylation changes in the brain and placenta have been published (41, 163, 164) and transcriptome analysis of placenta samples has been started with NGS. It will be interesting to see if new mutations from the PEG3 domain are found and related to IUGR and stillbirth. The genetic research behind gonadal hypoplasia in Northern Finncattle will continue, possibly revealing new aspects of cell migration during embryogenesis. Furthermore, the bovine PIRM model could represent the basis for comparative studies with mice and humans.

6 CONCLUDING REMARKS

During this PhD work we discovered two new inherited diseases in Ayrshire cattle and located causative mutations for these defects. The deletion in PEG3 domain cause IUGR and stillbirths and the mutation of *UBE3B* gene cause PIRM syndrome. Our studies also revealed an association between gonadal hypoplasia and the Cs_{29} allele in Northern Finncattle and Swedish Mountain cattle. The mutations related to these three defects or syndromes can now be easily tested for. The expression analyses proved that the mutations in Ayrshires also affected the RNA expression of the respective mutated gene. With the IUGR and stillbirths project we went a step further and demonstrated changed expression levels of neighbouring genes and hundreds of other genes in the foetal side of the placenta.

To reach these goals, this PhD study combined clinical veterinary medicine with basic genetic and molecular biology research. It has required successful and broad collaboration among veterinarians, molecular biologists and bioinformaticists, both nationally and internationally. This broad cooperation has met the demands of the specialised expertise from different fields and sophisticated facilities needed for this type of applied study. By utilizing pedigree, breeding and health records, and in combination with good animal management, this work has shown that Finnish cattle are an excellent subject for studying inherited diseases and causative mutation mapping. The phenotypes collected, multiple sample types, genotyping and sequence data will serve as a new infrastructure for Finnish bovine inheritance and basic biology studies.

It is likely that inbreeding of dairy cattle will lead to the emergence of new inherited defects with severe symptoms and adverse influence on functional traits. Furthermore, genetic causes for several traits are still unknown. Therefore, from the perspective of animal welfare and population fitness, and because of the economic importance of cattle, it is important to understand and contain these diseases. The effective control requires rapid detection of new defects and causative mutations. The results of this should be included in genomic testing of

breeding animals, especially the AI bulls. Lastly, information related to new defects and carriers among AI bulls needs to be made public. After all, the final breeding decisions are made by the farmers.

ACKNOWLEDGEMENTS

This thesis was carried out at the Department of Production Animal Medicine, Faculty of Veterinary Medicine, University of Helsinki. The research was supported by Academy of Finland, Finnish Veterinary Foundation, Niemi Foundation, Finnish Cultural Foundation, Finnish Foundation of Veterinary Research, Orion-Farmos Foundation, Emil Aaltonen Foundation, Elvi ja Eilo Eriksson Foundation and Betty Väänänen Foundation.

The thesis was made under the supervision of professors Magnus Andersson and Hannes Lohi. I owe my deepest gratitude to them for giving me an opportunity to be part of inspiring research environments and research projects. Both of my supervisors have offered me continual encouragement and willing to help for which I am truly thankful.

My gratitude is due to reviewers professor Jørgen S Agerholm and professor Göran Andersson for their critical evaluation of my thesis. For the skilful language revision, I would like to thank docent Jonathan Robinson. Professor Asko Mäki-Tanila is sincerely thanked for accepting the important post of an opponent at the dissertation.

Special thanks belong to Krzysztof Flisikowski for fruitful and rewarding collaboration in all projects included in the thesis. Thank you Krzysztof. Participation of Juhani Taponen especially in clinical analyses and sample collection is truly appreciated. The bioinformatic work and constructive criticism of Hubert Pausch has been indispensable. The generous help of Alexander Kind with the manuscript writing is acknowledged. I would also like to thank professor Angelika Schnieke for the opportunities to work in her laboratory in Freising. I really appreciate the friendly atmosphere and helpful attitude of the group. The same implies to the group of Ruedi Fries. I would like to express my thankfulness also to co-authors professors Marek Switonski, Jeremy F. Taylor, Susanne E. Ulbrich, Eckhard Wolf and Pekka Uimari as well as Dr. Joanna Nowacka-Woszuk, Dr. Stephanie D. McKay, Dr. Robert Schnabel, Dr. Hermann Schwarzenbacher, Dr. Izabela Szczerbal, Dr. Stefan Bauersachs, Dr.

Rainer W. Fürst, Dr. Anja Saalfrank, Dr. Michal Wysocki, Dr. Christine Wurmser, M.Sc. Reetta Hänninen and Antti Flyckt.

Special thanks belong to my co-workers in Saari unit and in Biomedicum. I have enjoyed to work with you and I appreciate yours expertise, encouragement and help. Especially warm thanks for co-PhD students Kaisa Kyöstilä, Leena Seppä-Lassila, Saija Ahonen, Pälvi Palander and Eeva Mustonen. Thank you professors Satu Pyörälä, Terttu Katila and Olli Peltoniemi for help and interest of my research career and Dr. Helena Rautala for all your work to improve my writings. Assistance and help of Sini Karjalainen, Ranja Eklund, Reijo Jokivuori, Hanna Petäjä, Tuija Mikkonen and Marja-Liisa Tasanko is very much appreciated. I am indebted to postdocs Eija Seppälä, Lotta Koskinen and Elisa Nevalainen for all their guidance.

I would also like to thank DVM Henri Simonen, Auli Himanen, Seppo Lintukangas, Hans Stålhammar, Søren Borchersen, Margareta Håård and Mia Sjögren from VikingGenetics and Sari Alhainen from Semex Finland Oy for providing samples. I am particularly grateful to everyone who has participated in sample collection, especially professors Antti Iivanainen, Stig Einarsson, Juha Kantanen and Heriberto Rodriguez-Martinez as well as Dr. Christine Kopp-Kuhlman, DVM Kalle Hakala, DVM Angelo Carluccio Neto, DVM Mari Niemi, DVM Vera Hakala, DVM Leena Pohjola, DVM Sami Savolainen, Annika Rikberg, Harri Toikka, Hannu Vainio, Miikka Depner, Hanna Perttula, Maritta Leppälä and Kirsti Nieminen, have been great help. Special thanks belong also to all cattle owners that took part in this study. The histology work of Kati Holmsten is appreciated. Thanks to Minna Toivonen, Anna Oksa and Seppo Niskanen for pedigree information

Sydämelliset kiitokset lapsuuden perheelleni ja lähisuvulle, joka loi minulle pohjan, jonka avulla olen pystynyt kohtaamaan elämän eri tilanteet. Erityiset kiitokset vanhemmilleni, jotka ovat tukeneet minua vuosien varrella. Suurimmat kiitokset kuuluvat puolisolleni Timo Suontaukselle joka on jakanut kanssani kaikki väitöskirjatyöhön liittyvät vaiheet. Kiitos jaksamisesta, välittämisestä ja avusta. Näiden vuosien varrella olet muokannut useita kuvia ja kommentoinut monia tekstejäni, josta olen kiitollinen. Apusi on ollut korvaamatonta.

REFERENCES

1. Utsunomiya YT, Perez O'Brien AM, Sonstegard TS, Solkner J, Garcia JF. Genomic data as the "hitchhiker's guide" to cattle adaptation: tracking the milestones of past selection in the bovine genome. Front Genet. 2015 Feb 10;6:36.

2. The state of the world's animal genetic resources for food and agriculture [Internet]. [cited 3/8/2015]. Available from: http://www.fao.org/docrep/010/a1250e/a1250e00.htm.

3. Aro J, Hilpelä-Lallukka R, Toivonen M, Vahlsten T. Mittaa ja valitse: Lypsykarjanjalostuksella tuloksiin. Tauriainen S, editor. Helsinki: Opetushallitus; 2012.

4. Bovine genomics. Womack JE, editor. Ames, Iowa: Wiley-Blackwell; 2012.

5. Nagahata H. Bovine leukocyte adhesion deficiency (BLAD): a review. J Vet Med Sci. 2004 Dec;66(12):1475-82.

6. OMIA - Online Mendelian Inheritance in Animals [Internet]. [cited 9/20/2014]. Available from: http://omia.angis.org.au/home/.

7. Drogemuller C, Reichart U, Seuberlich T, Oevermann A, Baumgartner M, Kuhni Boghenbor K, et al. An unusual splice defect in the mitofusin 2 gene (MFN2) is associated with degenerative axonopathy in Tyrolean Grey cattle. PLoS One. 2011 Apr 15;6(4):e18931.

8. Fasquelle C, Sartelet A, Li W, Dive M, Tamma N, Michaux C, et al. Balancing selection of a frame-shift mutation in the MRC2 gene accounts for the outbreak of the Crooked Tail Syndrome in Belgian Blue Cattle. PLoS Genet. 2009 Sep;5(9):e1000666.

9. Sartelet A, Klingbeil P, Franklin CK, Fasquelle C, Geron S, Isacke CM, et al. Allelic heterogeneity of Crooked Tail Syndrome: result of balancing selection? Anim Genet. 2012 Oct;43(5):604-7.

10. Kadri NK, Sahana G, Charlier C, Iso-Touru T, Guldbrandtsen B, Karim L, et al. A 660-Kb deletion with antagonistic effects on fertility and milk production segregates at high frequency in Nordic Red cattle: additional evidence for the common occurrence of balancing selection in livestock. PLoS Genet. 2014 Jan;10(1):e1004049.

11. Littlejohn MD, Henty KM, Tiplady K, Johnson T, Harland C, Lopdell T, et al. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nat Commun. 2014 Dec 18;5:5861.

12. Koch CT, Bruggmann R, Tetens J, Drogemuller C. A non-coding genomic duplication at the HMX1 locus is associated with crop ears in highland cattle. PLoS One. 2013 Oct 23;8(10):e77841.

13. Capitan A, Allais-Bonnet A, Pinton A, Marquant-Le Guienne B, Le Bourhis D, Grohs C, et al. A 3.7 Mb deletion encompassing ZEB2 causes a novel polled and multisystemic syndrome in the progeny of a somatic mosaic bull. PLoS One. 2012;7(11):e49084.

14. Cooper TA, Wiggans GR, Null DJ, Hutchison JL, Cole JB. Genomic evaluation, breed identification, and discovery of a haplotype affecting fertility for Ayrshire dairy cattle. J Dairy Sci. 2014 Mar 26;97(6):3878-82.

15. Bovine HapMap Consortium, Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science. 2009 Apr 24;324(5926):528-32.

16. Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science. 2009 Apr 24;324(5926):522-8.

17. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and characterization of a high density SNP genotyping assay for cattle. PLoS One. 2009;4(4):e5350.

18. Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, et al. Highly effective SNP-based association mapping and management of recessive defects in livestock. Nat Genet. 2008 Apr;40(4):449-54.

19. Pausch H, Kolle S, Wurmser C, Schwarzenbacher H, Emmerling R, Jansen S, et al. A nonsense mutation in TMEM95 encoding a nondescript transmembrane protein causes idiopathic male subfertility in cattle. PLoS Genet. 2014 Jan;10(1):e1004044.

20. Jung S, Pausch H, Langenmayer MC, Schwarzenbacher H, Majzoub-Altweck M, Gollnick NS, et al. A nonsense mutation in PLD4 is associated with a zinc deficiency-like syndrome in Fleckvieh cattle. BMC Genomics. 2014 Jul 22;15(1):623,2164-15-623.

21. Lehnert K, Ward H, Berry SD, Ankersmit-Udy A, Burrett A, Beattie EM, et al. Phenotypic population screen identifies a new mutation in bovine DGAT1 responsible for unsaturated milk fat. Sci Rep. 2015 Feb 26;5:8484.

22. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brondum RF, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat Genet. 2014 Aug;46(8):858-65.

23. Zhang MQ, Xu X, Luo SJ. The genetics of brown coat color and white spotting in domestic yaks (Bos grunniens). Anim Genet. 2014 Oct;45(5):652-9.

24. Durkin K, Coppieters W, Drogemuller C, Ahariz N, Cambisano N, Druet T, et al. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. Nature. 2012 Feb 1;482(7383):81-4.

25. Blake A, Pickford K, Greenaway S, Thomas S, Pickard A, Williamson CM, et al. MouseBook: an integrated portal of mouse resources. Nucleic Acids Res. 2010 Jan;38(Database issue):D593-9.

26. Geneimprint [Internet]. [cited 2/24/2015]. Available from: http://www.geneimprint.com/site/home.

27. Moore T, Haig D. Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet. 1991 Feb;7(2):45-9.

28. Renfree MB, Suzuki S, Kaneko-Ishino T. The origin and evolution of genomic imprinting and viviparity in mammals. Philos Trans R Soc Lond B Biol Sci. 2013 Jan 5;368(1609):20120151.

29. Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. Nat Rev Genet. 2011 Jul 18;12(8):565-75.

30. Adalsteinsson BT, Ferguson-Smith AC. Epigenetic control of the genome-lessons from genomic imprinting. Genes (Basel). 2014 Aug 14;5(3):635-55.

31. Koerner MV, Pauler FM, Huang R, Barlow DP. The function of non-coding RNAs in genomic imprinting. Development. 2009 Jun;136(11):1771-83.

32. Kim J, Bergmann A, Lucas S, Stone R, Stubbs L. Lineage-specific imprinting and evolution of the zinc-finger gene ZIM2. Genomics. 2004;84(1):47-58.

33. Choo JH, Kim JD, Kim J. Imprinting of an evolutionarily conserved antisense transcript gene APeg3. Gene. 2008 Feb 15;409(1-2):28-33.

34. Kim J, Bergmann A, Choo JH, Stubbs L. Genomic organization and imprinting of the Peg3 domain in bovine. Genomics. 2007;90(1):85-92.

35. Kim J, Ekram MB, Kim H, Faisal M, Frey WD, Huang JM, et al. Imprinting control region (ICR) of the Peg3 domain. Hum Mol Genet. 2012 Jun 15;21(12):2677-87.

36. Maegawa S, Itaba N, Otsuka S, Kamitani H, Watanabe T, Tahimic CG, et al. Coordinate downregulation of a novel imprinted transcript ITUP1 with PEG3 in glioma cell lines. DNA Res. 2004 Feb 29;11(1):37-49.

37. Kim JD, Kim J. YY1's longer DNA-binding motifs. Genomics. 2009 Feb;93(2):152-8.

38. Kim J, Kollhoff A, Bergmann A, Stubbs L. Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. Hum Mol Genet. 2003 Feb 1;12(3):233-45.

39. Thiaville MM, Kim H, Frey WD, Kim J. Identification of an evolutionarily conserved cis-regulatory element controlling the Peg3 imprinted domain. PLoS One. 2013 Sep 10;8(9):e75417.

40. Kim J, Noskov VN, Lu X, Bergmann A, Ren X, Warth T, et al. Discovery of a novel, paternally expressed ubiquitin-specific processing protease gene through comparative analysis of an imprinted region of mouse chromosome 7 and human chromosome 19q13.4. Genome Res. 2000 Aug;10(8):1138-47.

41. Venhoranta H, Li S, Salamon S, Flisikowska T, Andersson M, Switonski M, et al. Non-CpG hypermethylation in placenta of mutation-induced intrauterine growth restricted bovine foetuses. Biochem Biophys Res Commun. 2014 Feb 14;444(3):391-4.

42. He H, Kim J. Regulation and function of the peg3 imprinted domain. Genomics Inform. 2014 Sep;12(3):105-13.

43. Romagnoli S, Schlafer DH. Disorders of sexual differentiation in puppies and kittens: a diagnostic and clinical approach. Vet Clin North Am Small Anim Pract. 2006 May;36(3):573-606.

44. Palmieri C, Schiavi E, Della Salda L. Congenital and acquired pathology of ovary and tubular genital organs in ewes: a review. Theriogenology. 2011 Feb;75(3):393-410.

45. Abe S, Miyake YI, Kageyama SI, Watanabe G, Taya K, Kawakura K. Deletion of the Sry region on the Y chromosome detected in a case of equine gonadal hypoplasia (XY female) with abnormal hormonal profiles. Equine Vet J. 1999 Jul;31(4):336-8.

46. Makinen A, Hasegawa T, Makila M, Katila T. Infertility in two mares with XY and XXX sex chromosomes. Equine Vet J. 1999 Jul;31(4):346-9.

47. Bidarkar SS, Hutson JM. Evaluation and management of the abnormal gonad. Semin Pediatr Surg. 2005 May;14(2):118-23.

48. Akkoyunlu G, Tepekoy F, Bebis A, Uysal F. Bilateral total ovarian hypoplasia in a Holstein Friesian heifer. Acta Histochem. 2014 Oct;116(8):1519-21.

49. Vaughan EK, Long SE, Parkinson TJ, Smith KC, Noakes DE. Ovarian hypoplasia in Lleyn ewes. Vet Rec. 1997 Jan 25;140(4):100-1.

50. Purohit GN. Ovarian and oviductal pathologies in the buffalo: Occurrence, diagnostic and therapeutic approaches. Asian Pacific Journal Of Reproduction. 2014;3(2):156-68.

51. Settergren I. The ovarian morphology in clinical bovine gonadal hypoplasia with some aspects of its endogrine relations. Stockholm: Almqvist & Wiksells; 1964.

52. Eriksson K. Hereditary forms of sterility in cattle. Biological and genetical investigations. Lund: Håkan Ohlssons Boktryckeri; 1943.

53. Lagerlöf N, Settergren I. Gonadenhypoplasie beim Rind der schwedischen Gebirgsrasse. Zuchthygiene. 1961;5:141-58.

54. Lagerlöf N, Settergren I. Results of 17 years control of hereditary ovarian hypoplasia in cattle of the swedish highland breed. The Cornell veterinarian. 1953;43(1):52-64.

55. Lagerlöf N, Boyd H. Ovarian hypoplasia and other abnormal conditions in the sexual organs of cattle of the swedish highland breed - results of post-mortem examination of over 6,000 cows. The Cornell veterinarian. 1953;43(1):64-79.

56. Wriedt C. Colorsided cattle: Some remarks concerning their occurence and heredity. Journal of Heredity. 1925;16(2):51-6.

57. Olson T. Genetics of Colour Variation. In: Fries R, Ruvinsky A, editors. The Genetics of Cattle. Wallingford: Cabi publishing; 1999. p. 33-51.

58. Brenig B, Beck J, Floren C, Bornemann-Kolatzki K, Wiedemann I, Hennecke S, et al. Molecular genetics of coat colour variations in White Galloway and White Park cattle. Animal Genetics. 2013.

59. Grosz MD, MacNeil MD. The "spotted" locus maps to bovine chromosome 6 in a Hereford-Cross population. J Hered. 1999 Jan-Feb;90(1):233-6.

60. Fontanesi L, Tazzoli M, Russo V, Beever J. Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. Anim Genet. 2010 Jun;41(3):295-303.

61. Reinsch N, Thomsen H, Xu N, Brink M, Looft C, Kalm E, et al. A QTL for the degree of spotting in cattle shows synteny with the KIT locus on chromosome 6. J Hered. 1999 Nov-Dec;90(6):629-34.

62. Liu L, Harris B, Keehan M, Zhang Y. Genome scan of pigmentation traits in Friesian-Jersey crossbred cattle. J Genet Genomics. 2009 Nov;36(11):661-6.

63. David VA, Menotti-Raymond M, Wallace AC, Roelke M, Kehler J, Leighty R, et al. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. G3 (Bethesda). 2014 Aug 1;4(10):1881-91.

64. Cvejic D, Steinberg TA, Kent MS, Fischer A. Unilateral and bilateral congenital sensorineural deafness in client-owned pure-breed white cats. J Vet Intern Med. 2009 Mar-Apr;23(2):392-5.

65. OMIM - Online Mendelian Inheritance in Man [Internet]. [cited 3/2/2015]. Available from: http://omim.org/.

66. MGI-Mouse Genome Informatics-The international database resource for the laboratory mouse [Internet]. [cited 3/4/2015]. Available from: http://www.informatics.jax.org/.

67. Bernex F, De Sepulveda P, Kress C, Elbaz C, Delouis C, Panthier JJ. Spatial and temporal patterns of c-kit-expressing cells in WlacZ/+ and WlacZ/WlacZ mouse embryos. Development. 1996 Oct;122(10):3023-33.

68. Buehr M, McLaren A, Bartley A, Darling S. Proliferation and migration of primordial germ cells in We/We mouse embryos. Dev Dyn. 1993 Nov;198(3):182-9.

69. Nishikawa S, Kusakabe M, Yoshinaga K, Ogawa M, Hayashi S, Kunisada T, et al. In utero manipulation of coat color formation by a monoclonal anti-c-kit antibody: two distinct waves of c-kit-dependency during melanocyte development. EMBO J. 1991 Aug;10(8):2111-8.

70. Yoshida H, Kunisada T, Grimm T, Nishimura EK, Nishioka E, Nishikawa SI. Review: melanocyte migration and survival controlled by SCF/c-kit expression. J Investig Dermatol Symp Proc. 2001 Nov;6(1):1-5.

71. Tarbashevich K, Raz E. The nuts and bolts of germ-cell migration. Curr Opin Cell Biol. 2010 Dec;22(6):715-21.

72. Richardson BE, Lehmann R. Mechanisms guiding primordial germ cell migration: strategies from different organisms. Nat Rev Mol Cell Biol. 2010 Jan;11(1):37-49.

73. Aoki H, Yamada Y, Hara A, Kunisada T. Two distinct types of mouse melanocyte: differential signaling requirement for the maintenance of non-cutaneous and dermal versus epidermal melanocytes. Development. 2009 Aug;136(15):2511-21.

74. Ogawa M, Nishikawa S, Yoshinaga K, Hayashi S, Kunisada T, Nakao J, et al. Expression and function of c-Kit in fetal hemopoietic progenitor cells: transition from the early c-Kit-independent to the late c-Kit-dependent wave of hemopoiesis in the murine embryo. Development. 1993 Mar;117(3):1089-98.

75. Pedurupillay CR, Baroy T, Holmgren A, Blomhoff A, Vigeland MD, Sheng Y, et al. Kaufman oculocerebrofacial syndrome in sisters

with novel compound heterozygous mutation in UBE3B. Am J Med Genet A. 2015 Mar;167(3):657-63.

76. Basel-Vanagaite L, Dallapiccola B, Ramirez-Solis R, Segref A, Thiele H, Edwards A, et al. Deficiency for the ubiquitin ligase UBE3B in a blepharophimosis-ptosis-intellectual-disability syndrome. Am J Hum Genet. 2012 Dec 7;91(6):998-1010.

77. Basel-Vanagaite L, Yilmaz R, Tang S, Reuter MS, Rahner N, Grange DK, et al. Expanding the clinical and mutational spectrum of Kaufman oculocerebrofacial syndrome with biallelic UBE3B mutations. Hum Genet. 2014 Mar 11.

78. Flex E. Loss of function of the E3 ubiquitin-protein ligase UBE3B causes Kaufman oculocerebrofacial syndrome. Journal of Medical Genetics, 17 May 2013. 2013.

79. Froyen G, Belet S, Martinez F, Santos-Reboucas CB, Declercq M, Verbeeck J, et al. Copy-number gains of HUWE1 due to replication- and recombination-based rearrangements. Am J Hum Genet. 2012 Aug 10;91(2):252-64.

80. Niemeyer CM, Kang MW, Shin DH, Furlan I, Erlacher M, Bunin NJ, et al. Germline CBL mutations cause developmental abnormalities and predispose to juvenile myelomonocytic leukemia. Nat Genet. 2010 Sep;42(9):794-800.

81. Martinelli S, De Luca A, Stellacci E, Rossi C, Checquolo S, Lepri F, et al. Heterozygous germline mutations in the CBL tumor-suppressor gene cause a Noonan syndrome-like phenotype. Am J Hum Genet. 2010 Aug 13;87(2):250-7.

82. Ohta A, Schumacher FR, Mehellou Y, Johnson C, Knebel A, Macartney TJ, et al. The CUL3-KLHL3 E3 ligase complex mutated in Gordon's hypertension syndrome interacts with and ubiquitylates WNK isoforms: disease-causing mutations in KLHL3 and WNK4 disrupt interaction. Biochem J. 2013 Apr 1;451(1):111-22.

83. Fang P, Lev-Lehman E, Tsai TF, Matsuura T, Benton CS, Sutcliffe JS, et al. The spectrum of mutations in UBE3A causing Angelman syndrome. Hum Mol Genet. 1999 Jan;8(1):129-35.

84. Matsuura T, Sutcliffe JS, Fang P, Galjaard RJ, Jiang YH, Benton CS, et al. De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. Nat Genet. 1997 Jan;15(1):74-7.

85. Scheffner M, Kumar S. Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects. Biochim Biophys Acta. 2014 Jan;1843(1):61-74.

86. Gong TW, Huang L, Warner SJ, Lomax MI. Characterization of the human UBE3B gene: structure, expression, evolution, and alternative splicing. Genomics. 2003 Aug;82(2):143-52.

87. Locke MM, Penedo MC, Bricker SJ, Millon LV, Murray JD. Linkage of the grey coat colour locus to microsatellites on horse chromosome 25. Anim Genet. 2002 Oct;33(5):329-37.

88. Cribiu EP, Di Berardino D, Di Meo GP, Eggen A, Gallagher DS, Gustavsson I, et al. International System for Chromosome Nomenclature of Domestic Bovids (ISCNDB 2000). Cytogenet Cell Genet. 2001;92(3-4):283-99.

89. BAC Clones distribution center - BACPAC Resources Center [Internet]. [cited 7/8/2014]. Available from: http://bacpac.chori.org/home.htm.

90. Switonski M, Gustavsson I, Ploen L. The nature of the 1;29 translocation in cattle as revealed by synaptonemal complex analysis using electron microscopy. Cytogenet Cell Genet. 1987;44(2-3):103-11.

91. Switonski M, Andersson M, Nowacka-Woszuk J, Szczerbal I, Sosnowski J, Kopp C, et al. Identification of a new reciprocal translocation in an AI bull by synaptonemal complex analysis, followed by chromosome painting. Cytogenet Genome Res. 2008;121(3-4):245-8.

92. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 2009;10(4):R42.

93. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559-75.

94. Seaton G, Hernandez-Sanchez J, Grunchec JA, White I, Allen J, De Koning DJ, et al. GridQTL: A Grid Portal for QTL Mapping of Compute Intensive Datasets. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production (Belo horizonte, Brazil), ; 2006.

95. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999 Dec;55(4):997-1004.

96. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009 Feb;84(2):210-23.

97. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 2007 Nov;17(11):1665-74.

98. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010 Mar 1;26(5):589-95.

99. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078-9.

100. Picard [Internet]. [cited 6/11/2014]. Available from: http://picard.sourceforge.net/.

101. Florea L, Souvorov A, Kalbfleisch TS, Salzberg SL. Genome assembly has a major impact on gene content: a comparison of

annotation in two Bos taurus assemblies. PLoS One. 2011;6(6):e21400.

102. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000;132:365-86.

103. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 1998 Mar;8(3):175-85.

104. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998 Mar;8(3):186-94.

105. Nickerson DA, Tobe VO, Taylor SL. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Res. 1997 Jul 15;25(14):2745-51.

106. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res. 1998 Mar;8(3):195-202.

107. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000 Sep 8;302(1):205-17.

108. Weksberg R, Hughes S, Moldovan L, Bassett AS, Chow EW, Squire JA. A method for accurate detection of genomic microdeletions using real-time quantitative PCR. BMC Genomics. 2005 Dec 13;6:180.

109. Lachman HM, Pedrosa E, Petruolo OA, Cockerham M, Papolos A, Novak T, et al. Increase in GSK3beta gene copy number variation in bipolar disorder. Am J Med Genet B Neuropsychiatr Genet. 2007 Apr 5;144B(3):259-65.

110. Furst RW, Kliem H, Meyer HH, Ulbrich SE. A differentially methylated single CpG-site is correlated with estrogen receptor alpha transcription. J Steroid Biochem Mol Biol. 2012 May;130(1-2):96-104.

111. Gremme G. Engineering a software tool for gene structure prediction in higher organisms. Information And Software Technology, 2005, Vol 47(15), pp 965-978. 2005;47(15):965-78.

112. Quackenbush J, Cho J, Lee D, Liang F, Holt I, Karamycheva S, et al. The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. Nucleic Acids Res. 2001 Jan 1;29(1):159-64.

113. Lewis SE, Searle SM, Harris N, Gibson M, Lyer V, Richter J, et al. Apollo: a sequence annotation editor. Genome Biol. 2002;3(12):RESEARCH0082.

114. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res. 2003 Jul 1;31(13):3568-71.

115. Huber W, von Heydebreck A, Sultmann H, Poustka A, Vingron M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics. 2002;18 Suppl 1:S96-104.

116. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. 2004;3:Article3.

117. ClustalW2 [Internet]. [cited 6/11/2014]. Available from: http://www.ebi.ac.uk/Tools/msa/clustalw2/.

118. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc. 2009;4(3):363-71.

119. Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, et al. Characterization of a common susceptibility locus for asthma-related traits. Science. 2004 Apr 9;304(5668):300-4.

120. Lauvergne JJ. Gonadal hypoplasia and white coat color in Swedish Highland cattle. J Hered. 1970 Jan-Feb;61(1):43-4.

121. Jansen S, Aigner B, Pausch H, Wysocki M, Eck S, Benet-Pages A, et al. Assessment of the genomic variation in a cattle population by re-sequencing of key animals at low to medium coverage. BMC Genomics. 2013 Jul 4;14:446,2164-14-446.

122. Kim J, Frey WD, He H, Kim H, Ekram MB, Bakshi A, et al. Peg3 mutational effects on reproduction and placenta-specific gene families. PLoS One. 2013 Dec 31;8(12):e83359.

123. Curley JP, Barton S, Surani A, Keverne EB. Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. Proc Biol Sci. 2004 Jun 22;271(1545):1303-9.

124. Swaney WT, Curley JP, Champagne FA, Keverne EB. Genomic imprinting mediates sexual experience-dependent olfactory learning in male mice. Proc Natl Acad Sci U S A. 2007 Apr 3;104(14):6084-9.

125. Li L, Keverne EB, Aparicio SA, Ishino F, Barton SC, Surani MA. Regulation of maternal behavior and offspring growth by paternally expressed Peg3. Science. 1999 Apr 9;284(5412):330-3.

126. Wei W, Pelechano V, Jarvelin AI, Steinmetz LM. Functional consequences of bidirectional promoters. Trends Genet. 2011 Jul;27(7):267-76.

127. Moore GE, Ishida M, Demetriou C, Al-Olabi L, Leon LJ, Thomas AC, et al. The role and interaction of imprinted genes in human fetal growth. Philos Trans R Soc Lond B Biol Sci. 2015 Mar 5;370(1663):20140074.

128. Tunster SJ, Jensen AB, John RM. Imprinted genes in mouse placental development and the regulation of fetal energy stores. Reproduction. 2013 Apr 29;145(5):R117-37.

129. Piedrahita JA. The role of imprinted genes in fetal growth abnormalities. Birth Defects Res A Clin Mol Teratol. 2011 Aug;91(8):682-92.

130. Fowden AL, Coan PM, Angiolini E, Burton GJ, Constancia M. Imprinted genes and the epigenetic regulation of placental phenotype. Prog Biophys Mol Biol. 2011 Jul;106(1):281-8.

131. Lefebvre L, Viville S, Barton SC, Ishino F, Keverne EB, Surani MA. Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene Mest. Nat Genet. 1998 Oct;20(2):163-9.

132. Gerber S, Vardhana S, Meagher-Villemure K, Vial Y, Hohlfeld P, Witkin SS. Association between fetal interleukin-1 receptor antagonist gene polymorphism and unexplained fetal death. Am J Obstet Gynecol. 2005 Oct;193(4):1472-7.

133. Amash A, Holcberg G, Sapir O, Huleihel M. Placental secretion of interleukin-1 and interleukin-1 receptor antagonist in preeclampsia: effect of magnesium sulfate. J Interferon Cytokine Res. 2012 Sep;32(9):432-41.

134. Clark EA, Mele L, Wapner RJ, Spong CY, Sorokin Y, Peaceman A, et al. Association of fetal inflammation and coagulation pathway gene polymorphisms with neurodevelopmental delay at age 2 years. Am J Obstet Gynecol. 2010 Jul;203(1):83.e1,83.e10.

135. Hefler LA, Reyes CA, O'Brien WE, Gregg AR. Perinatal development of endothelial nitric oxide synthase-deficient mice. Biol Reprod. 2001 Feb;64(2):666-73.

136. Barut F, Barut A, Gun BD, Kandemir NO, Harma MI, Harma M, et al. Intrauterine growth restriction and placental angiogenesis. Diagn Pathol. 2010 Apr 22;5:24,1596-5-24.

137. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015 Jan 23;347(6220):1260419.

138. Ilmarinen P, James A, Moilanen E, Pulkkinen V, Daham K, Saarelainen S, et al. Enhanced expression of neuropeptide S (NPS) receptor in eosinophils from severe asthmatics and subjects with total IgE above 100IU/ml. Peptides. 2014 Jan;51:100-9.

139. Vendelin J, Pulkkinen V, Rehn M, Pirskanen A, Raisanen-Sokolowski A, Laitinen A, et al. Characterization of GPRA, a novel G protein-coupled receptor related to asthma. Am J Respir Cell Mol Biol. 2005 Sep;33(3):262-70.

140. Kumsta R, Chen FS, Pape HC, Heinrichs M. Neuropeptide S receptor gene is associated with cortisol responses to social stress in humans. Biol Psychol. 2013 May;93(2):304-7.

141. Banister CE, Koestler DC, Maccani MA, Padbury JF, Houseman EA, Marsit CJ. Infant growth restriction is associated with distinct patterns of DNA methylation in human placentas. Epigenetics. 2011 Jul;6(7):920-7.

142. Koukoura O, Sifakis S, Spandidos DA. DNA methylation in the human placenta and fetal growth (review). Mol Med Rep. 2012 Apr;5(4):883-9.

143. Kay GW, Grobbelaar JA, Hattingh J. Heritable testicular hypoplasia in Nguni (Bos indicus) bulls: vascular characteristics and testosterone production. J Reprod Fertil. 1992 Nov;96(2):537-47.

144. Pretorius A, Osbourn DE. Observations on hypoplasia in Nguni cattle. Die Afrikanerbees joernaal = The Afrikaner cattle journal. 1979;25(1):40-7.

145. Kay GW, Meyer EHH. Testicular compensation in Nguni (Bos indicus, Sanga) bulls with unilateral gonadal hypoplasia and aplasia. South African Journal of Animal Science. 1985;15(2):58-60.

146. Chahrour MH, Yu TW, Lim ET, Ataman B, Coulter ME, Hill RS, et al. Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. PLoS Genet. 2012;8(4):e1002635.

147. Semerci CN, Kalay E, Yıldırım C, Dincer T, Olmez A, Toraman B, et al. Novel splice-site and missense mutations in the ALDH1A3 gene underlying autosomal recessive anophthalmia/microphthalmia. Br J Ophthalmol. 2014 Feb 25.

148. Du H, Sheriff S, Bezerra J, Leonova T, Grabowski GA. Molecular and enzymatic analyses of lysosomal acid lipase in cholesteryl ester storage disease. Mol Genet Metab. 1998 Jun;64(2):126-34.

149. Terracina M, Posteraro P, Schubert M, Sonego G, Atzori F, Zambruno G, et al. Compound heterozygosity for a recessive glycine substitution and a splice site mutation in the COL7A1 gene causes an unusually mild form of localized recessive dystrophic epidermolysis bullosa. J Invest Dermatol. 1998 Nov;111(5):744-50.

150. Lomax MI, Huang L, Cho Y, Gong TL, Altschuler RA. Differential display and gene arrays to examine auditory plasticity. Hear Res. 2000 Sep;147(1-2):293-302.

151. The Human Protein Atlas [Internet]. [cited 3/6/2015]. Available from: http://www.proteinatlas.org/.

152. Mileski A. Expression Pattern of Ubiquitin Ligase E3 (UBE3B) in Mouse. 35th Annual Meeting and Exhibition of the American Association for Dental Research, 30th Annual Meeting of the Canadian Association for Dental Research and 83rd Annual Meeting and Exhibition of American Dental Education Association. 2006.

153. Braganza A. Understanding the Function of UBE3B and ANKZF1 in Response to Mitochondrial Stress. Environmental And Molecular Mutagenesis, 2013, Vol 54, pp S31-S31. 2013;54:S31-.

154. Yanase S, Ishi N. Cloning of the oxidative stress-responsive genes in Caenorhabditis elegans. J Radiat Res. 1999 Mar;40(1):39-47.

155. Asselbergs FW, Guo Y, van Iperen EP, Sivapalaratnam S, Tragante V, Lanktree MB, et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. Am J Hum Genet. 2012 Nov 2;91(5):823-38.

156. Porter FD, Herman GE. Malformation syndromes caused by disorders of cholesterol synthesis. J Lipid Res. 2011 Jan;52(1):6-34.

157. Knürr T, Mäntysaari E. Genominen data tuo hyviä ja huonoja uutisia. Nauta. 2014;5:12,12-13.

158. Washburn SP, Silvia WJ, Brown CH, McDaniel BT, McAllister AJ. Trends in reproductive performance in Southeastern Holstein and Jersey DHI herds. J Dairy Sci. 2002 Jan;85(1):244-51.

159. Lucy MC. Reproductive loss in high-producing dairy cattle: where will it end? J Dairy Sci. 2001 Jun;84(6):1277-93.

160. Haplotype tests for recessive disorders that affect fertility and other traits [Internet].; 2015 [cited 3/16/2015]. Available from: http://aipl.arsusda.gov/reference/recessive_haplotypes_ARR-G3.html.

161. Magee DA, Spillane C, Berkowicz EW, Sikora KM, MacHugh DE. Imprinted loci in domestic livestock species as epigenomic targets for artificial selection of complex traits. Anim Genet. 2014 Aug;45 Suppl 1:25-39.

162. Magee DA, Sikora KM, Berkowicz EW, Berry DP, Howard DJ, Mullen MP, et al. DNA sequence polymorphisms in a panel of eight candidate bovine imprinted genes and their association with performance traits in Irish Holstein-Friesian cattle. BMC Genet. 2010 Oct 13;11:93,2156-11-93.

163. Venhoranta H, Bauersachs S, Taponen J, Lohi H, Taira T, Andersson M, et al. Fetal growth restriction caused by MIMT1 deletion alters brain transcriptome in cattle. Int J Dev Neurosci. 2013 May 29;31(7):463-7.

164. Flisikowski K, Venhoranta H, Bauersachs S, Hanninen R, Furst RW, Saalfrank A, et al. Truncation of MIMT1 Gene in the PEG3 Domain Leads to Major Changes in Placental Gene Expression and Stillbirth in Cattle. Biol Reprod. 2012 Oct 24;87(6).